(2)

Validation Test Report for the First-Generation Dart Gulf Stream Forecasting System

Theodore J. Bennett, Jr.
Ocean Sensing and Prediction Division
Ocean Science Directorate

Robert C. Rhodes
Naval Oceanographic Office
Stennis Space Center, MS 39529

Foreword

Activities at sea, ranging from commercial fishing to acoustic surveillance, are influenced by the mesoscale structure of the ocean. Considerable spatial and temporal variability due to meandering fronts and drifting rings characterize the mesoscale structure. Knowledge of the mesoscale structure can make activities at sea more cost effective.

Recognizing the need to map and understand the mesoscale structure of the ocean, the Naval Oceanographic and Atmospheric Research Laboratory, in conjunction with the U.S. Naval Oceanographic Office, is developing and transitioning a system for making nowcasts and forecasts of the Gulf Stream frontal path. This report presents the results of experiments that estimate the performance of the system in an operational, rather than research-grade, environment. In addition, operator guidance is provided for using this system.

William B Moseley

W. B. Moseley
Technical Director

L.R. Elliott

L. R. Elliott, Commander, USN
Officer in Charge

Executive Summary

Acoustic propagation within the Northwest Atlantic is highly dependent on the meandering Gulf Stream front and its associated drifting rings, the so-called mesoscale structure of the ocean. Considerable variability in space and time characterize this structure.

This report documents the validation of a system that makes nowcasts and forecasts of the Gulf Stream frontal path. The validation was done as part of the transition of the first-generation Data Assimilation Research and Transition (DART) model to operational Navy use. Two sets of experiments designed to evaluate the forecast skill of the system in a quasi-operational environment are the crux of the validation. In one set of experiments, operational front and ring maps are used to initialize and verify 1- and 2-week forecast experiments. In the other set of experiments, the value the DART model adds to the present operational capability is estimated. The proposed operational use of the model is simulated by setting up a simple data assimilation system that uses a 1-week forecast to interpolate across long data gaps that arise during the mapping of front and ring positions. The gappy paths are constructed by superimposing observed data gaps on complete frontal paths. This report also provides guidance to an operator on how to use a model forecast to fill in the data gaps.

Both sets of experiments show the system to be a statistically significant improvement over an assumption of persistence, i.e., no change over time. The principal technical issues identified in the report are the inference of the Gulf Stream north wall from the forecasted axis path, as well as the need to develop and test a more objective means to assimilate observations and a forecast into a composite frontal path. It is recommended that these issues be pursued.

Acknowledgments

LCDR William Cook, Program Manager at the Space and Naval Warfare Systems Command, funded this transition at NOARL through the Navy Ocean Modeling and Prediction (NOMP) Program, which is managed by Mr. Robert Peloquin. The Program Element is 0603207N. Within NOMP, this transition is part of the Data Assimilation Research and Transition (DART) project. We gratefully acknowledge Mr. Dan Fox for providing the statistical analysis of the results. We extend our thanks and appreciation to the numerous people at the Naval Oceanographic and Atmospheric Research Laboratory (NOARL) and the U.S. Naval Oceanographic Office (NAVOCEANO) that have contributed software, guidance, and other assistance to this work. In particular, we thank Dr. Michael Carnes, Mr. Dan Fox, Dr. John Harding, Dr. Joseph McCaffrey, and Dr. Jim Mitchell (DART Manager) at NOARL and Dr. Charles Horton and Mr. Andrew Johnson at NAVOCEANO. We also acknowledge the contributions of the members of the Technical Validation Panel, Mr. R. Michael Clancy, Dr. John Harding, and Dr. John Hovermale (Chair) of NOARL and Dr. William Schmitz of Woods Hole Oceanographic Institution. The work on graphics software and software documentation by Ms. Julia Crout, Ms. Barbara Ray, and Mr. Michael Rich, of Planning Systems, Inc., is appreciated.

The mention of commercial products or the use of company names does not in any way imply endorsement by the U.S. Navy or NOARL.

Contents

Validation Test Report for the First-Generation Dart Gulf Stream Forecasting System

I. Introduction

A number of activities at sea, ranging from commercial fishing to acoustic surveillance, are influenced by the mesoscale thermal structure of the ocean. Recognizing the importance of the mesoscale structure of the ocean, the Data Assimilation Research and Transition (DART) project of the Naval Oceanographic and Atmospheric Research Laboratory (NOARL) is developing and transitioning to operational Navy use a system for making nowcasts and forecasts of the Gulf Stream. In its initial implementation, the objective of the system is to make nowcasts, as well as 7- and 14-day forecasts, of the path of the Gulf Stream and its interactions with associated warm-core and cold-core rings.

In December 1989, the DART project presented to the Commander, Naval Oceanography Command (CNOC) Independent Model Review Panel (CIMREP), the results of a proof-of-concept study of the first-generation DART system. The key data set for this study was a set of maps of Gulf Stream frontal paths and ring locations that were obtained during periods of unusually abundant infrared (IR) data, as well as in situ and altimetry data. These data were sufficient to confidently map on a weekly basis the path and ring locations over virtually the entire domain from 78°W to 50°W for several periods of 3- to 5-week duration. In the study, the DART model was initialized from each of several of these maps and the forecast verified against the map that was valid at the time of forecast verification. The results demonstrated that the DART 7- and 14-day forecasts of the Gulf Stream frontal path are better than persistence, which assumes that the frontal path is unchanged from the initial state.

Several issues arise when the forecast system is considered for use in a more operationally realistic mode. In the operational world, clouds typically mask a significant fraction of the Gulf Stream path, thereby creating data gaps in the forecast initialization. In addition, the operationally available front and ring maps do not necessarily have the same quality as the research-grade maps prepared during periods of unusually abundant IR data. The effect on forecast skill of clouds and the relative utility of operational vs. research-grade maps must be evaluated as part of the transition of the system to operational use.

With the signing of a Memorandum of Understanding (CNOC ltr, Ser 3140, 14 February 1990) a formal transition process has been established between NOARL, U.S. Naval Oceanographic Office (NAVOCEANO) and the Fleet Numerical Oceanography Center (FNOC). A key aspect of the memorandum is a transition plan that identifies some of the technical issues to be addressed by the transition process, as well as the approach to be taken (see appendix).

The results of several types of experiments designed to evaluate the model in a more operationally realistic mode are presented here. In one set of experiments, a simple data assimilation system is set up by using a 1-week forecast to interpolate across long data gaps. The gappy paths are constructed by superimposing observed data gaps on complete frontal paths. This data assimilation system is compared to the existing operational capability. In another set of experiments, forecasts are made using the front and ring maps prepared operationally by NAVOCEANO to initialize the model. In addition, the potential role that an operator would have in the use of this system as an aid to the construction of a front and ring map is described.

The report is structured as follows. Sections II and III present the components of the forecast system and some technical aspects of the system. Section IV presents the results of the proof-of-concept study. Section V provides results from experiments using operational, rather than research-grade, paths for initialization and verification. Section VI compares the forecast system to the present operational capability. Section VII compares pairs of forecasts that differ by small, localized differences in their initial states and examines the evolution of these differences. Section VIII describes the potential role of an operator while using this system as an aid in constructing a front and ring map. Section IX gives a summary and the conclusions, and Section X gives recommendations.

II. Forecast System Components

The Gulf Stream forecast capability is made up of several components. The system is initialized from a front and ring map, hereafter referred to as a "bogus." The bogus is constructed by hand using a disparate set of frontal position data. It can also be constructed via an operator/machine mix using the operational NAVOCEANO product PATHFINDER, which determines a frontal path from a disparate set of frontal position data via optimal interpolation. The bogus is then used to construct a three-dimensional analysis of temperature and salinity using the Optimal Thermal Interpolation System (OTIS) and its associated feature models. The analysis field from OTIS is then used to calculate a dynamic height field, which in turn is used to initialize the upper layer of the forecast model. The output from the circulation model is a set of predicted frontal paths that can be used as data for later bogus preparations.

A. Frontal Bogus

The lack of sufficient data at depth in the ocean makes it difficult to obtain a three-dimensional analysis of temperature, salinity, and sound velocity of the ocean environment. The most abundant oceanic data are available at the surface, namely satellite IR data, altimetry, drifting buoys, and multichannel sea surface temperatures. An important operational product of the Operational Oceanography Center (OOC) at NAVOCEANO is the surface front and ring bogus which uses surface data to obtain an analysis of the position of the north wall of the Gulf Stream and positions and sizes of warm and cold rings.

Currently, satellite IR images that depict the surface temperature of the North Atlantic region are the most abundant data sources used in analyzing the Gulf Stream path and rings. The bogus is prepared three times weekly using the previous bogus, which is combined with the most recent data to obtain a best-guess composite that depicts the positions on that day.

The most obvious problem is that clouds limit the visibility of the Gulf Stream and create data gaps where the operator has no information on the position of the features. Another problem is that the surface temperature analysis is not always representative of the actual location of features. Misrepresentation is especially a problem with rings as they lose their surface signature with age. Also, the surface north wall position of the Gulf Stream is always not reliable for depicting the actual path of the maximum transport of the Gulf Stream. These problems are not easily overcome, but satellite altimetry shows potential for eliminating some of these problems. However, no altimetry data are operationally available at this time.

B. NAVOCEANO PATHFINDER Module

As discussed earlier, limited data coverage is the main problem in preparation of the bogus. At NAVOCEANO, an operational optimal interpolation scheme, PATHFINDER, is used for blending different data types of various ages into a composite analysis of the location of the north wall of the Gulf Stream. PATHFINDER converts latitude/longitude frontal positions to cross-stream/downstream positions relative to a defined mean path and assigns an uncertainty to each observation based on its age. It then performs a one-dimensional optimal interpolation of cross-stream position versus downstream position. The objective analysis procedure is taken from Bretherton et al. (1976). When the procedure is completed, PATHFINDER converts cross-stream/downstream positions back to latitude/longitude coordinates.

In the assimilation approach taken in this report, PATHFINDER will be an integral part of the process for using the model forecast as data across gaps in the analysis. The 7- and 14-day forecasts of the model can be input into PATHFINDER as data, provided an estimated error is also input. PATHFINDER will then do a composite analysis using a blend of observed data and model forecast data. Some preliminary results of this technique, as well as the direction of further development, will be discussed in section VIII.

C. OTIS Analysis

The OTIS system has been used operationally at FNOC for several years (Clancy et al., 1990). It is operational as a regional product in the Gulf Stream using in situ data, frontal bogus, and water-mass-based climatology as input. The bogus is an integral part of the analysis scheme as it defines the positions of the features that are depicted in the three-dimensional analysis using the feature models within OTIS. The method of using OTIS to construct three-dimensional temperature and salinity fields from only the front and ring bogus is explained in Fox et al. (1990).

In all of the experiments discussed here, the only data available for OTIS were the front and ring boguses. Although OTIS can assimilate data and altimetry data, this capability was not used because the means to use these data are not currently operational at the OOC. The purpose of these experiments is to simulate operational capabilities at NAVOCEANO. In situ data and altimetry were used to locate front and ring features in the paths used for the proof-of-concept study. However, the operational boguses are based only on the IR data that are operationally analyzed at the OOC.

D. Ocean Circulation Module

The OTIS three-dimensional temperature and salinity analysis is used to construct a nowcast dynamic height field relative to 1000 decibars. This field is converted to an upper layer pressure field (p_1) for NOARL's two-layer, limited-area, Gulf Stream circulation model. The basic circulation model is discussed by Hurlburt and Thompson (1980) and Wallcraft (1990), and its application to the Gulf Stream is described by Thompson and Schmitz (1989). The circulation model initialization uses the nowcast pressure field for the upper layer and a statistically predicted lower layer pressure field (p_2). The statistical prediction is based on regression coefficients of the relationship between upper and lower layer pressures that were obtained from many years of model integration (see Hurlburt et al., 1990). The model and the initialization techniques are discussed in detail by Fox et al. (1990).

The model produces a 14-day forecast of the Gulf Stream and associated rings. Because of the lack of large-scale forcing in the limited area model, ring motions are not realistic and are not verified. The axis of the Gulf Stream is output from the initialization and for the 7- and 14- day forecast positions. The current operational Gulf Stream forecast model at NAVOCEANO runs on a 7-day cycle, and the DART model would initially be used in the same manner. The forecasts of the frontal path will be used to fill in data gaps in a simple data assimilation technique that should improve the bogus quality. The quality of forecasts that are subsequently initialized from these boguses should also be improved.

III. Technical Aspects of the DART System

A. Estimating the Gulf Stream North Wall from the Frontal Axis

The DART forecast system outputs a prediction of the path of the Gulf Stream frontal axis. The OOC, however, depicts the path of the north wall of the Gulf Stream. Hence, the forecasted axis path must be converted into a path of the north wall if it is to be useful in constructing the next bogus using PATHFINDER. It should be noted that the frontal axis is determined from the pressure field, and the north wall is observed by examining thermal gradients at the surface; hence, these paths are quite distinct entities.

A new technique has been developed for use with the operational boguses. The decision was made to do this conversion in a manner that is consistent with the techniques already in place at the OOC. In its standard usage, the PATHFINDER program outputs a simulated axis of the Gulf Stream if it is given the north wall position. To output this information, it uses a frontal feature model that estimates the offset between the north wall and the axis. The effect of curvature is included in the feature model. PATHFINDER also has the capability to go the other way and output the north wall if it is given the axis. In practice, this output is accomplished by asking PATHFINDER to estimate the path of the frontal axis at a certain negative depth. This technique has been incorporated into the system. All of the results obtained using the operational boguses that are presented here use north wall paths determined by PATHFINDER from the forecast axis.

Figures 1a, 1b, and 1c indicate the results of the PATHFINDER technique. The solid line to the north in each figure is the original north wall path from the OOC that was used to initialize the DART system on 2 May 1990, 9 May 1990, and 16 May 1990, respectively. The solid line to the south is the axis output obtained using the DART forecast initialization process. The dashed line to the north is the north wall obtained from PATHFINDER using the DART axis as input. The results show that the

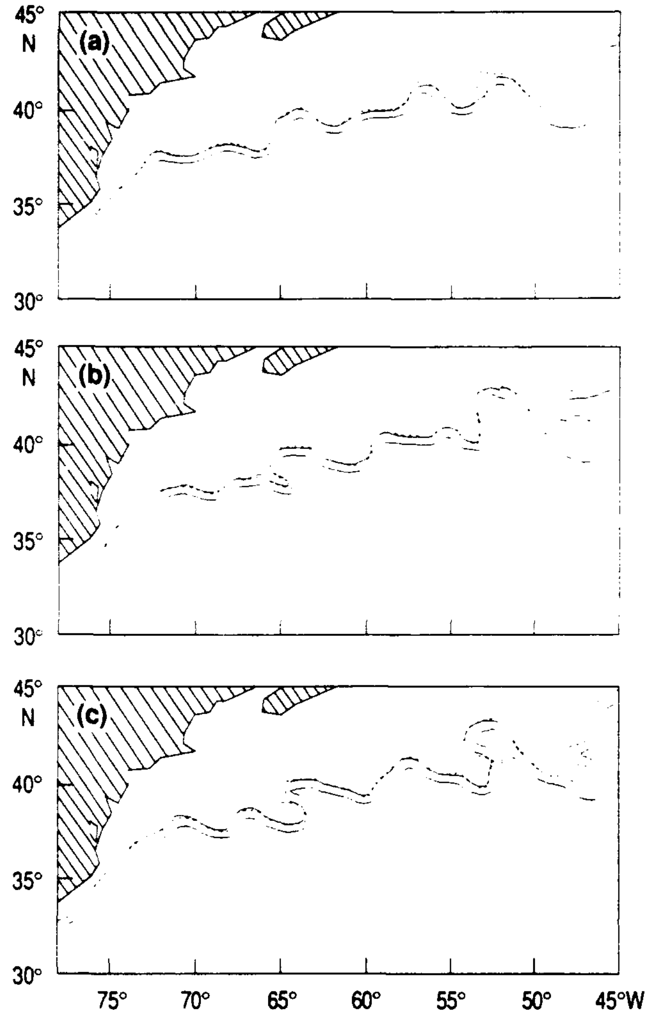

Figure 1. Three comparisons of the operational north wall bogus prepared by the OOC (northern solid line), the inferred frontal axis of the model initialization generated from the OOC bogus (southern solid line) and the path of the north wall inferred from the frontal axis using PATHFINDER (dashed line) for year/Julian day. (a) 90/122, (b) 90/129, and (c) 90/136.

technique matches well the original OOC bogus. The largest differences are seen in tight meanders as would be expected. The mean difference between the inferred and OOC north walls is only about 9 km. This error source is small compared to the other error sources, but it needs to be examined in further studies in order to develop improved techniques that will minimize the error.

B. Model Ports

This section describes two versions of the DART model. These versions were tested during the evaluation phase. The advantages and disadvantages of each and the reasons for making the final choice are presented. The two versions differ in the way the inflow port information is specified. For short (2 week) forecasts, the forecast skill is not significantly altered by the choice of port method, so the version that was easiest to implement in the operational environment was selected.

To compute the statistical regression coefficients required to initialize the forecast system, the circulation model had to be run for approximately 20 model-years. During this run, the upper and lower layer inflows (that is, the Florida Current and the Deep Western Boundary Current) are specified as boundary conditions that are fixed at their climatological positions and strengths. These locations and amplitudes produce a Gulf Stream whose mean axis closely matches that seen in data, and whose variability (both in magnitude and location) also matches that seen in data (both from GEOSAT and in situ measurements). Changes to these boundary conditions can result in significant changes in the mean path and variability of the Gulf Stream (Thompson et al., 1989).

During the period leading up to the CIMREP evaluation in December 1989, it was felt that maintaining proper boundary conditions was an important consideration in initializing forecasts. Since the weekly Gulf Stream axis was often not coincident with its climatological position (used in the circulation model), it was therefore necessary to alter the axis in the far western domain to gradually taper from the position provided by the bogus to the position required by the model. All the forecasts that were run for the CIMREP evaluation were performed using this sort of initialization.

An alternate method for initializing the circulation model was developed more than a year prior to the CIMREP meeting for the first version of the DART forecast breadboard (Fox et al., 1988). This technique is referred to as the "modified" or "adaptive" port version. In this method, rather than supplying the inflow based on climatological values and locations, the initial pressure fields (p_1 and inferred p_2) are used to calculate the implied geostrophic current profiles at each of the circulation model inflow ports. The ports are substantially widened so that nearly any location of the Florida Current can be handled. Since the ports are implemented as model boundary conditions, the flow implied in the initial state is held constant for the entire forecast. That is, the inflow is not permitted to vary either in magnitude or location.

The main advantage of the adaptive port version of the system is that it automatically adapts to nearly any location of the Gulf Stream axis. One disadvantage of this method is that, by changing the boundary conditions for the circulation model, the mean path and variability pattern will almost certainly also be changed. One problem is potentially more serious: with the change in the location and strength of the inflow ports, the statistical inference coefficients employed to provide the initial lower layer pressure field will be incorrect near the ports. The first problem is serious only for long runs of the model; the second could affect the forecast skill immediately.

To evaluate the difference in the two port versions, forecasts were made using both the original CIMREP data sets (for which the axis of the stream had been altered to properly match the climatological axis location used by the circulation model at the western inflow port) and the independent data set from May to July 1989 (for which no such modifications to the axis were done). Table 1 shows that in neither case was the forecast skill of the system drastically affected. The rows labelled "Difference" are calculated by subtracting the forecast error made by the fixed-port version from the error made by the adaptive-port version. Positive differences imply that the adaptive version is "worse" than the nonadaptive version. Using the CIMREP data (with the altered axis locations), the forecast made by the adaptive version of the system is insignificantly different at 1 week and is generally less than 1 km worse at 2 weeks, compared to the nonadaptive version. Using the newer boguses (with unaltered axes), the forecasts are insignificantly different at 1 week, and at 2 weeks the adaptive port version is a very slight improvement (less than 0.5 km) over the nonadaptive version. The last result is particularly interesting, since in nonadaptive forecasts where the location of the stream has moved significantly from its climatological position, the mismatch between the model port and the stream in the data induces the generation of substantial Kelvin waves near the ports. The forecast is visually unsatisfactory near these boundaries, but the evolution of the axis itself is not significantly affected.

Table 1. Comparison of forecasts made with the adaptive and fixed port versions of the DART system. The rows labelled "Difference" are calculated by subtracting the forecast error made by the fixed-port version from the error made by the adaptive-port version. Positive differences imply that the adaptive version is "worse" than the nonadaptive version.

ORIGINAL CIMREP DATASETS—FORECAST RESULTS					
		ERROR (km OFFSET) BY REGION			
PORT VERSION	DURATION (weeks)	ENTIRE AREA	WESTERN	CENTRAL	EASTERN
Fixed	1	33.2 ± 4.7	24.3 ± 4.5	28.1 ± 5.4	48.0 ± 11.3
Adaptive	1	33.3 ± 4.7	24.3 ± –4.4	28.0 ± 5.5	48.2 ± 11.1
Difference	1	0.06 ± 0.08	–0.02 ± 0.18	0.16 ± 0.13	0.02 ± 0.28
Fixed	2	37.6 ± 3.7	25.4 ± 2.1	39.5 ± 6.7	45.9 ± 7.4
Adaptive	2	38.4 ± 3.6	25.9 ± 2.1	40.8 ± 6.2	46.4 ± 7.2
Difference	2	0.78 ± 0.13	0.53 ± 0.29	1.28 ± 0.66	0.45 ± 0.43
NEW MAY–JULY 1989 DATASETS—FORECAST RESULTS					
		ERROR (km OFFSET) BY REGION			
PORT VERSION	DURATION (weeks)	ENTIRE AREA	WESTERN	CENTRAL	EASTERN
Fixed	1	37.7 ± 2.8	26.9 ± 3.1	50.0 ± 7.4	38.7 ± 6.1
Adaptive	1	37.8 ± 2.8	27.1 ± 3.2	50.0 ± 7.4	38.6 ± 6.3
Difference	1	0.05 ± 0.09	0.18 ± 0.13	0.02 ± 0.16	–0.15 ± 0.27
Fixed	2	47.8 ± 2.6	37.6 ± 4.8	62.6 ± 7.7	45.4 ± 4.1
Adaptive	2	47.6 ± 2.6	37.7 ± 4.9	62.0 ± 7.7	45.3 ± 4.2
Difference	2	–0.25 ± 0.06	0.05 ± 0.19	–0.63 ± 0.43	–0.15 ± 0.18

Since the operational forecasts will not exceed 2 weeks, the longer term problems inherent in the adaptive port version will not be apparent, but this version has an advantage: it can accept the bogus without modification of the location of the stream at the western boundary of the domain. Given these operational conditions, and based on the results in Table 1, the adaptive port version of the DART system was determined to be more appropriate for transition.

IV. Summary of Proof-of-Concept Results

The DART project was requested to perform a proof-of-concept study by the CIMREP panel in the fall of 1989. The experiments were designed to investigate the current research capability within DART for forecasting the path of the Gulf Stream. In particular, the DART capability would be compared with the baseline operational capability at NAVOCEANO in a set of parallel experiments. The baseline capability, the Navy Operational Gulf Stream Forecast System (NOGUFS) 1.0, is based on the Harvard GulfCast system. It was evaluated, and the results indicated that the quality of NOGUFS 1.0 forecasts was about equal to that of a persistence forecast at 7 days and worse at 14 days. The operational capability needed to be upgraded to a system that was better than a persistence forecast.

At the same time that NOGUFS 1.0 was being evaluated operationally at NAVOCEANO, a NOARL research project was concluding its own evaluation of NOGUFS 1.0. An important product of NOARL's evaluation was a set of research quality boguses that were prepared by making use of abundant IR, air-dropped expendable bathythermographs from field experiments, and GEOSAT altimetry data. The boguses were carefully constructed and quality control checked to form the best available maps for forecast validation. These boguses were used as input data for the parallel comparison between NOGUFS 1.0 and the DART model. These paths are referred to as the CIMREP paths and are also used in some of the experiments presented here.

Figure 2 compares the results of the validation tests of several versions of NOGUFS and the DART forecast system. The results are depicted as a average absolute offset error of the 7- and 14-day forecast frontal position error of the two systems for all of the hindcast cases. The results clearly indicate that DART outperforms NOGUFS 1.0 and is also better than a persistence forecast of both 7 and 14 days. These results convinced the CIMREP panel to recommend transition of the DART system to NAVOCEANO. The complete discussion of the evaluation techniques and the results can be found in Fox et al. (1990).

V. Experiments with Operational Boguses

A. Results

All of the previous validation tests performed with the DART model were done using research-quality bogus data sets that had more abundant data coverage than is normally observed in the Gulf Stream. An important test criterion was how the DART system would perform using the real-time operational front and ring product from the OOC. A test was performed using approximately 15 weekly cases. The model was initialized with the OOC's operational front and ring composite, and forecasts were performed for 14 days with the weekly boguses used to validate the 7- and 14-day forecasts. An example of an initialization, a 7-day forecast, and the validation are shown in Figures 3a, 3b, and 3c, respectively. The tests began on 90/122 (May 2, 1990), with a 7-day analysis

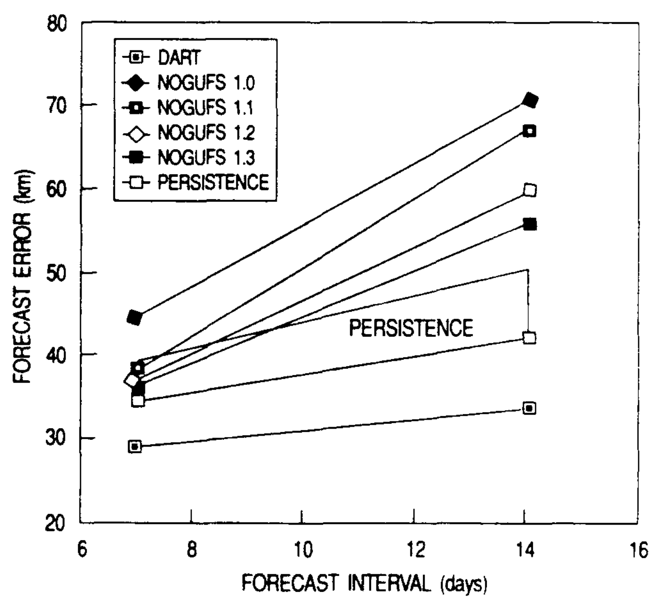

Figure 2. Comparison of the average absolute offset error of forecasts made using the DART model, various versions of the NOGUFS 1.x model and persistence at 7 and 14 days. There were abundant data for initialization and verification of these forecasts.

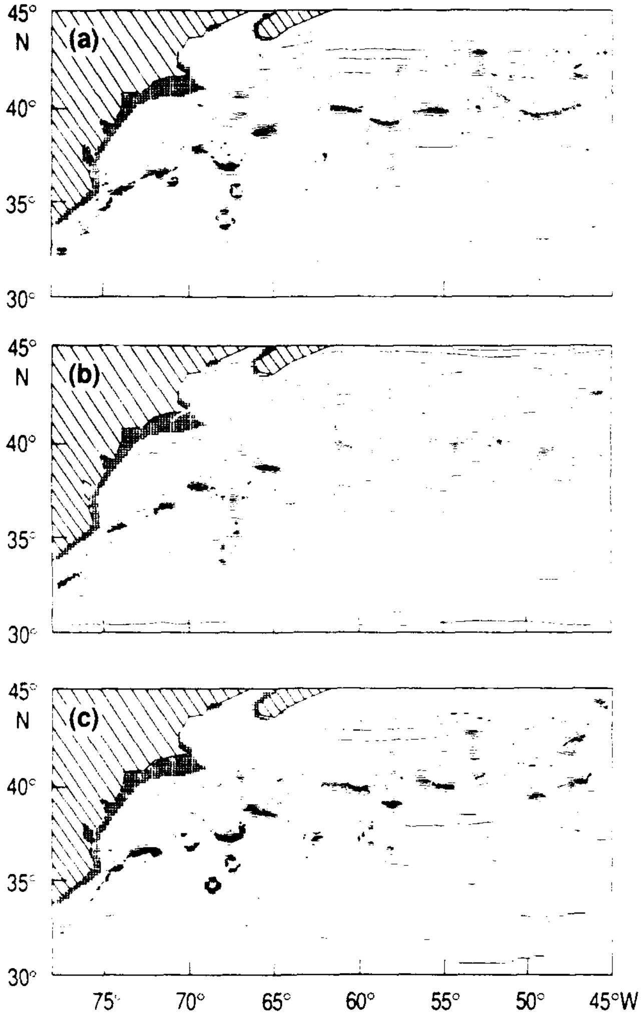

Figure 3. The dynamic height field from (a) the initialization of Julian Day 157, 1990, (b) the 1-week forecast valid Julian Day 164, 1990, and (c) the verification dated Julian Day 164, 1990.

interval, and continued using data through 90/234 (August 22, 1990). The model is only being validated from 73° to 53°W. These domains are the same domains used in the evaluation of the Harvard GulfCast system.

Table 2 shows the results for the 7-day validation from Julian days 122–227. The results are shown for the entire area from 73° to 53°W, as well as the western (73°–66°W), central (66°–59°W) and eastern (59°–53°W) sections in order to examine forecast skill as a function of area. The error estimates (km) are obtained by examining the average absolute offset between the forecast and bogus paths. An asterisk indicates that the DART forecast error was less than the persistence forecast error.

Figure 4 is a scatter plot of the data in Table 2 and indicates that for most of the cases, the forecast error is smaller than persistence error when compared to the OOC bogus of 1 week later. Also, when the model forecast error is greater than persistence, the difference between them is generally smaller than when the forecast beats persistence. The mean of all the data combined shows that the model beats persistence by about 10% at 7 days. Figure 5 shows the results in a histogram format with the number of cases within a certain error range depicted. The DART forecast is skewed more toward the lower error numbers than is persistence. Table 3 shows mean values broken down by area and indicates that the forecast beats persistence in all areas, with the largest difference in the central area. This difference is not surprising because the central area is the most likely for events, and also shows the value of the model in predicting ring formation and absorptions.

Table 4 shows similar results for the 14-day forecast evaluation. The forecast error is lower than persistence in all but 10 of 60 cases. The results indicate that the model is better relative to persistence at 14 days than at 7 days. Figure 6 is the scatter plot of the data and further indicates the model's superiority over persistence. In this case, the mean shows the model is beating persistence by 15%. The histogram, Figure 7, clearly indicates that the DART forecast is skewed toward lower error relative to persistence. Table 5 shows the means for all areas, which indicates that the model is superior in all sections and, like the 7-day results, the difference is greatest in the central area.

The result that the forecast is better relative to persistence at 14 days is understandable if the error associated with the bogus data is considered. The main data source for the bogus is IR data of the surface position of the north wall of the Gulf Stream, which has an error of at least 15 km in locating the stream. The lateral movement of the Gulf Stream in 7 days is estimated from operational bogus data to be 15–20 km (Fox et al., 1990). Because the movement at 7 days is small compared to the error of the bogus, the difference between forecast and persistence error is not as evident from the error statistics. At 14 days the lateral movement is 30 and 40 km, so the difference stands out from the error and is more clearly shown in the statistics. The purpose of the study using operational bogus data was to show that the model forecast was a better predictor than using persistence, which is currently used operationally at the OOC. Even though the error is large in positioning the Gulf Stream,

Table 2. Errors for ensemble of 7-day forecasts initialized and verified with operational boguses (F) compared to that of persistence (P).

JULIAN DAY	ENTIRE AREA 73°–53°W		WESTERN AREA 73°–66°W		CENTRAL AREA 66°–59°W		EASTERN AREA 59°–53°W	
	P	F	P	F	P	F	P	F
122	28.7	26.1*	28.0	20.4*	20.9	21.1	37.8	36.0*
129	39.3	30.7*	19.5	14.1*	46.3	35.3*	51.9	43.3*
136	32.8	39.2	12.4	24.7	62.5	66.2	22.8	25.3
143	44.9	46.9	11.9	16.1	81.3	77.7	40.3	46.7
150	33.2	33.7	38.5	31.5*	30.1	33.3	29.7	37.5
157	20.0	17.6*	21.0	21.1	19.6	18.0*	18.9	11.9*
164	35.9	31.7*	33.9	29.3*	37.2	34.1*	35.1	29.1*
171	37.9	21.8*	10.4	10.9	70.0	28.5*	24.2	24.0*
178	34.4	31.5*	24.0	16.9*	52.9	50.1*	18.6	19.5
185	29.1	22.7*	21.0	13.1*	31.2	23.5*	32.7	29.3*
192	35.7	28.4*	17.3	16.5*	23.5	14.6*	67.3	57.0*
199	68.9	49.4*	10.2	8.8*	80.4	35.9*	123.9	114.0*
206	26.5	24.0*	9.1	9.0*	29.5	30.5	42.4	33.4*
213	68.1	62.5*	15.5	9.5*	93.1	94.2	100.2	87.1*
220	26.5	27.6	23.6	27.6	29.5	29.4*	25.6	24.4*
227	19.4	19.0*	16.2	14.0	26.7	24.9*	10.7	15.1

*Forecast Error < Persistence Error

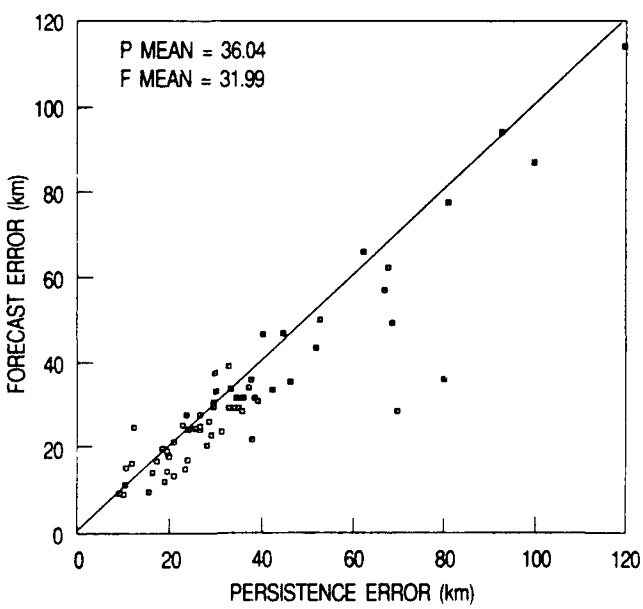

Figure 4. A scatter plot of the 7-day forecast error against the 7-day persistence error for a series experiments during Julian days 122–227, 1990. Operational boguses used for initialization and verification.

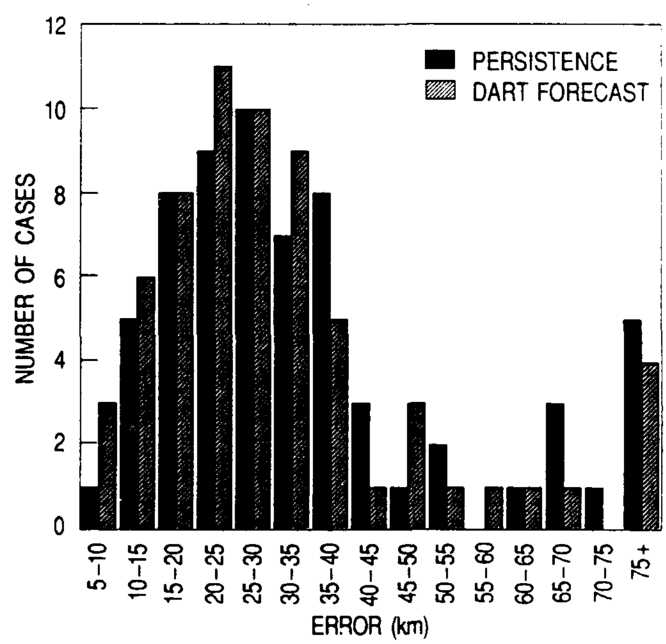

Figure 5. A histogram of the distribution of forecast and persistence errors for an ensemble of 7-day forecasts that used operational boguses for initialization and verification.

Table 3. Mean error for ensemble of 7-day forecasts initialized and verified with operational boguses compared to that of persistence for the entire domain and several subsections of the domain.

OPERATIONAL BOGUS DATA J. D. 122–227, 1990 7-DAY MEAN ERROR (km)				
	ENTIRE AREA 73°–53°W	WESTERN AREA 73°–66°W	CENTRAL AREA 66°–59°W	EASTERN AREA 59°–53°W
Persistence	36.33	19.53	45.91	42.63
DART Forecast	32.05	17.72	38.58	39.60

Table 4. Errors for ensemble of 14-day forecasts initialized and verified with operational boguses (F) compared to that of persistence (P) for the entire domain and several subsections of the domain.

DART FORECAST SYSTEM 14-DAY FORECAST								
	ENTIRE AREA 73°–53°W		WESTERN AREA 73°–66°W		CENTRAL AREA 66°–59°W		EASTERN AREA 59°–53°W	
JULIAN DAY	P	F	P	F	P	F	P	F
122	41.8	33.0*	40.1	22.0*	42.2	32.7*	42.6	47.2
129	51.6	42.0*	23.9	17.4*	73.9	69.7*	57.2	39.3*
136	52.4	47.0*	21.4	26.7	78.9	63.9*	57.6	51.6*
143	51.8	48.1*	40.2	33.2*	54.2	50.0*	65.7	67.6
150	29.1	30.8	35.2	26.4*	31.9	40.8	16.4	20.2
157	36.5	27.5*	25.4	13.0*	40.2	35.6*	41.4	27.7*
164	49.3	40.8*	38.0	29.1*	60.6*	59.0*	45.5	28.2*
171	46.8	39.8*	26.2	16.2*	70.9	60.5*	31.5	32.0
178	47.0	40.1*	25.4	20.5*	69.1	56.1*	36.0	35.4*
185	46.2	39.4*	23.2	19.5*	49.4	38.2*	60.7	57.4*
192	63.4	57.7*	22.1	20.5*	103.7	78.4*	62.7	75.2
199	73.6	49.7*	12.2	15.4	101.4	53.6*	108.3	81.6*
206	77.1	71.1*	16.7	9.1*	78.9	81.7	148.6	133.5*
213	54.8	48.1*	23.0	19.4*	61.3	53.6*	83.5	74.5*
220	23.3	19.0*	15.1	20.2	31.7	22.9*	19.5	9.5*

*Forecast Error < Persistence Error

the results indicate that the model outperforms persistence. All of the results shown used the entire OOC bogus as the Gulf Stream truth data to evaluate how the model would perform in a realistic operational environment.

An attempt was made to discard the cases where either the initialization bogus or the validation bogus was not as accurate because of a lack of quality IR data coverage. Table 6 shows the periods from Table 2 where the bogus quality was judged to be inferior. These cases are shown by NA where data was not available. The results still indicate that the model outperforms persistence when the questionable cases are thrown out. Figure 8 shows a histogram of the quality data cases and Table 7 the means for each area. The results continue to indicate model forecast superiority and show that for the central area, the model performance is better relative to persistence than in the case where all of the data was considered. Table 8, Figure 9, and Table 9 show similar results for the 14-day forecasts. These results also show model forecast superiority, and the statistics are similar to the results with all the data combined.

The quality data cases were chosen to give confidence in the operational results and to remove

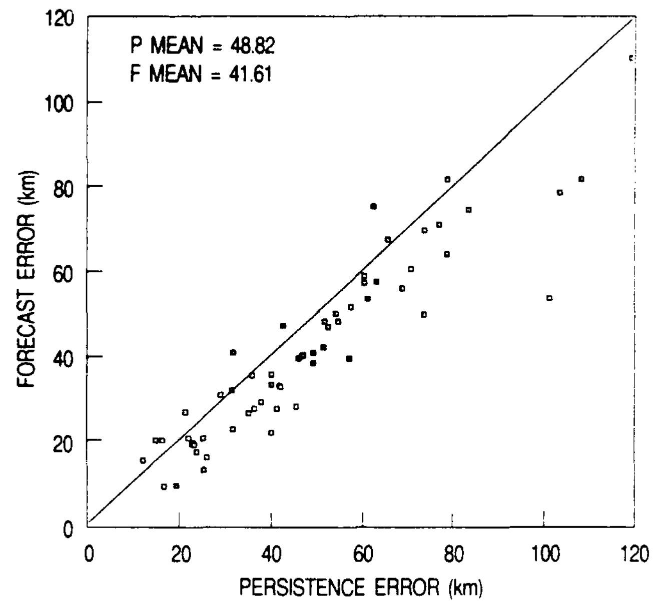

Figure 6. A scatter plot of the 14-day forecast error against the 14-day persistence error for a series experiments that used operational boguses for initialization and verification.

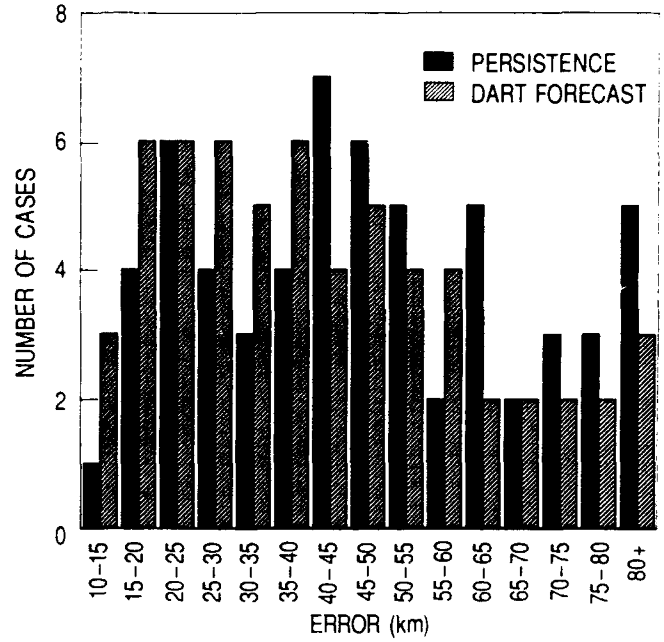

Figure 7. A histogram of the distribution of forecast and persistence errors for an ensemble of 14-day forecasts that used operational boguses for initialization and verification.

Table 5. Mean error for ensemble of 14-day forecasts initialized and verified with operational boguses compared to that of persistence for the entire domain and several subsections of the domain.

OPERATIONAL BOGUS DATA J. D. 122–220, 1990 14-DAY MEAN ERROR (km)				
	ENTIRE AREA 73°–53°W	WESTERN AREA 73°–66°W	CENTRAL AREA 66°–59°W	EASTERN AREA 59°–53°W
Persistence	49.64	25.87	63.22	58.48
DART Forecast	42.27	20.57	53.11	52.06

questions about the reliability of the operational tests, given the uncertainty of the weekly boguses. Because the results were similar to those cases in which all the data were considered, we are confident that the tests using the entire operational bogus are representative of the performance of the DART forecast system.

B. Statistical Significance

What is the statistical significance of the improvement of the forecast error over the persistence error? The significance is determined by how well the mean can be estimated, not by the distribution. As the sample becomes larger, the distribution remains the same, but the standard error of the mean decreases. Comparing the mean (error difference) to the standard error of the mean gives an estimate of the statistical significance.

Table 10 presents the results of a statistical analysis. The column labeled mean is the mean difference between the forecast error and the persistence error. A positive number indicates that the mean forecast error is smaller than the mean persistence error. The second column is an estimate of the standard error of the mean. The third column is an estimate of the statistical significance of the results. Two statistics can be used to estimate the significance level. If the quantity being analyzed is Gaussian, then the usual "t-test" can be made. The differences for each region were examined using the Shapiro-Wilk test for normality, which indicated that the distributions were usually not sufficiently

Table 6. Errors for ensemble of high-quality, 7-day forecasts initialized and verified with operational boguses (F) compared to that of persistence (P) for the entire domain and several subsections of the domain. High-quality forecasts are a subsample of Table 3. Low-quality forecasts are indicated by NA.

DART FORECAST SYSTEM 7-DAY FORECAST						
	WESTERN AREA 73°–66°W		CENTRAL AREA 66°–59°W		EASTERN AREA 59°–53°W	
JULIAN DAY	P	F	P	F	P	F
122	28.0	20.4*	20.9	21.1	NA	NA
129	19.5	14.1*	46.3	35.3*	51.9	43.3*
136	NA	NA	NA	NA	22.8	25.3
143	NA	NA	NA	NA	40.3	46.7
150	38.5	31.5*	NA	NA	29.7	37.5
157	NA	NA	NA	NA	NA	NA
164	NA	NA	NA	NA	NA	NA
171	10.4	10.9	70.0	28.5*	24.2	24.0*
178	24.0	16.9*	52.9	50.1*	18.6	19.5
185	21.0	13.1*	31.2	23.5*	32.7	29.3*
192	17.3	16.5*	23.5	14.6*	NA	NA
199	10.2	8.8*	80.4	35.9*	123.9	114.0*
206	9.1	9.0*	NA	NA	NA	NA
213	15.5	9.5*	NA	NA	NA	NA
220	23.6	27.6	29.5	29.4*	25.6	24.4*
227	NA	NA	NA	NA	NA	NA

*Forecast Error < Persistence Error

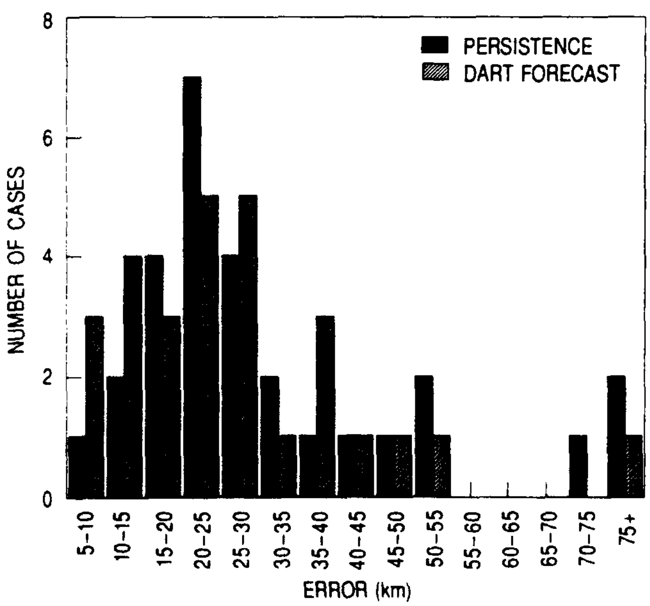

Figure 8. A histogram of the distribution of forecast and persistence errors for an ensemble of high-quality, 7-day forecasts that used operational boguses for initialization and verification.

Table 7. Mean error for ensemble of high-quality, 7-day forecasts initialized and verified with operational boguses compared to that of persistence for the entire domain and several subsections of the domain.

OPERATIONAL BOGUS DATA J. D. 122–227, 1990 7-DAY QUALITY DATA CASES MEAN ERROR (km)			
	WESTERN AREA 73°–66°W	CENTRAL AREA 66°–59°W	EASTERN AREA 59°–53°W
Persistence	19.73	44.34	41.08
DART Forecast	16.21	29.80	40.44

Table 8. Errors for ensemble of high-quality, 14-day forecasts initialized and verified with operational boguses (F) compared to that of persistence (P) for the entire domain and several subsections of the domain. High-quality forecasts are a subsample of Table 5. Low-quality forecasts are indicated by NA.

DART FORECAST SYSTEM 14-DAY FORECAST						
	WESTERN AREA 73°–66°W		CENTRAL AREA 66°–59°W		EASTERN AREA 59°–53°W	
JULIAN DAY	P	F	P	F	P	F
122	40.1	22.0*	42.2	32.7*	42.6	47.2
129	NA	NA	NA	NA	57.2	39.3*
136	21.4	26.7	NA	NA	57.6	51.6*
143	NA	NA	NA	NA	65.7	67.6
150	35.2	26.4*	NA	NA	16.4	20.2
157	25.4	13.0*	40.2	35.6*	41.4	27.7*
164	NA	NA	NA	NA	NA	NA
171	26.2	16.2*	70.9	60.5*	31.5	32.0
178	25.4	20.5*	69.1	56.1*	NA	NA
185	23.2	19.5*	49.4	38.2*	60.7	57.4
192	22.1	20.5*	103.7	78.4*	NA	NA
199	12.2	15.4	NA	NA	NA	NA
206	16.7	9.1*	78.9	81.7	148.6	133.5*
213	23.0	19.4*	NA	NA	NA	NA
220	NA	NA	NA	NA	NA	NA

*Forecast Error < Persistence Error

Gaussian. Hence, a nonparametric test for statistical significance is used for all regions and both weeks.

Consider the 1-week forecasts in the eastern domain as an example. The difference between the forecast error and the persistence error has a mean

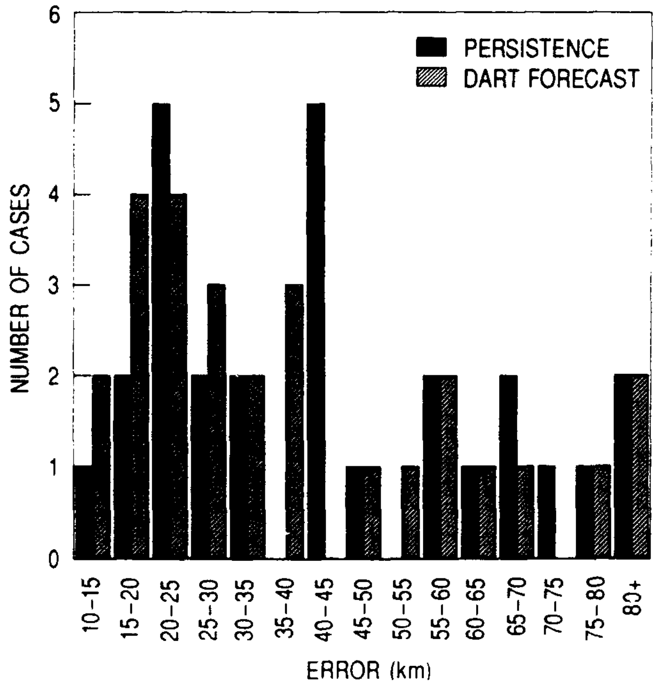

Figure 9. A histogram of the distribution of forecast and persistence errors for an ensemble of high-quality, 14-day forecasts that used operational boguses for initialization and verification.

Table 9. Mean error for ensemble of high-quality, 14-day forecasts initialized and verified with operational boguses compared to that of persistence for the entire domain and several subsections of the domain.

OPERATIONAL BOGUS DATA J. D. 122–220, 1990 14-DAY QUALITY DATA CASES MEAN ERROR (km)			
	WESTERN AREA 73°–66°W	CENTRAL AREA 66°–59°W	EASTERN AREA 59°–53°W
Persistence	24.63	64.91	57.97
DART Forecast	18.97	54.74	52.94

of 3.0 km, which is clearly greater than the standard error in the mean of 1.6 km. Hence, the mean is clearly different from zero. The chance that the true mean is actually zero is only 7.7%. In other words, we can be 92% confident that the mean is not zero and the result that the forecast model beats persistence is statistically significant. The results show that, for any region, we can be quite confident—92% or greater—that the forecast composite is better than the persistence composite at 1 week and at least 95% confident at 2 weeks.

Table 10. Results of statistical analysis using the operational boguses for initialization and verification. The mean of the difference in error (persistence error minus forecast error), the standard error (Stderr) of the mean, the statistical significance and the confidence in the results are shown for the entire area as well as the western, central, and eastern areas. A nonparametric distribution is assumed since the distribution is not sufficiently Gaussian. The results indicate that our confidence that the forecast has a smaller error than persistence ranges from 92% in the eastern domain at 1 week to over 99% at 2 weeks in the western and central domains.

	STATISTICAL ANALYSIS AT 1 WEEK			
REGION	MEAN (km)	STANDARD ERROR (km)	SIGNIFICANCE	CONFIDENCE (%)
Western	1.8	1.4	0.1046	90
Central	7.3	3.6	0.0615	94
Eastern	3.0	1.6	0.0773	92
Entire Area	4.3	1.6	0.0131	99
	STATISTICAL ANALYSIS AT 2 WEEKS			
REGION	MEAN (km)	STANDARD ERROR (km)	SIGNIFICANCE	CONFIDENCE (%)
Western	6.5	1.7	0.0084	99+
Central	10.1	3.4	0.0024	99+
Eastern	6.4	2.7	0.0479	95
Entire Area	7.4	1.4	0.0001	99+

VI. Comparison of the Forecast System to the Existing Operational Capability

A series of experiments are made to evaluate the DART system in a quasi-operational mode. To link these experiments to the operational world, observed data gaps are superimposed on the CIMREP paths used for model initialization. Previous 1-week forecasts are used as data across long data gaps in the initialization path, thereby simulating the operational use of the system. An operator has to make several judgments to run this system in such a data assimilation mode. These judgments are an integral part of the experiments presented in this section. The product of this system, which is a composite of frontal path data and a previous model forecast, is evaluated by comparing differences between the composite and the gapless CIMREP path to the differences between the CIMREP path and a composite of data and a persistence forecast. The latter composite mimics the operational capability. An estimate of the value added by the forecast system can be obtained from this comparison.

A. Observed Data Gaps

The observed data gaps are obtained from data supplied by the OOC. These data consist of 30, almost continuous weekly maps of the frontal path. Each map shows the path segments seen over the previous week or so, as well as the composite of these data. The analysis of the path segments and the composite were done by the OOC. Maps of data gaps are obtained by identifying the longitudinal segments of each map that have no data of age 7 days or younger.

The distribution in time and space of the age of the data is presented in Figure 10. Longitudinal bands with data of age 1–3 days, 4–7 days and greater than 7 days are shown. A band with data of age 4–7 days implies that there are no data of age 1–3 days. Bands with no data younger than 8 days are considered data gaps. The fraction of the path seen from July to November is typically greater than 80%. In December and January, the fraction is only 30%–70%. The gaps are also rather persistent in winter. The quantity of data is insufficient to determine how representative these data are of the long-term average.

B. Data Assimilation System

1. Description of Experiments

A simple data assimilation system is set up by using a 1-week forecast as data for interpolating across gaps in the frontal path data at the time of the forecast validation. Such a system brings data forward in time. Without the forecast model, the frontal path could only be determined from the previous week's data because the skill of persistence is small at much beyond 1 week. The forecast model is initialized from data that are 1 to 2 weeks old. The value of the forecast system is that it can be used to fill a data gap—not by assuming that the aged data is persistent, but by using an estimate of how the aged data has evolved in time.

The observations and the forecast path need not agree at the edges of the data gaps. In practice, the forecast is used as data across data gaps longer than about 2.5 degrees of longitude. The match between data and the forecast at the edge of the gap is judged to be too sensitive to latitudinal shifts and longitudinal phase shifts of the forecast path for the forecast to be directly substituted into a short gap. In these cases, PATHFINDER is used alone to interpolate across the data gap. This

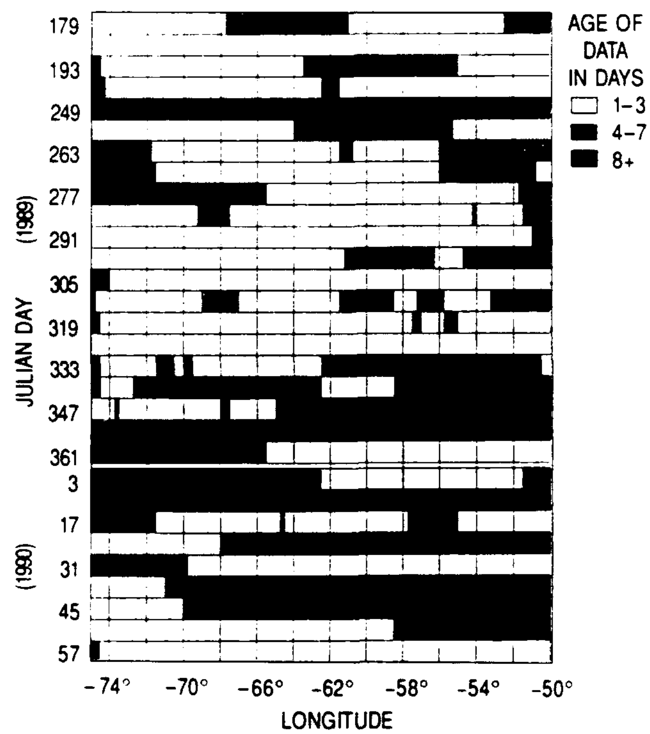

Figure 10. A plot of the age of the most recent IR data available to map the Gulf Stream frontal path as a function of longitude for each bogus in a nearly continuous series.

judgment is, of course, subjective, and is open to further examination.

Even in the case of a long data gap, there can be a disconnect between the observations and the forecast at the edge of the gap. The approach taken in these experiments is that the operator must reconcile the disconnect between the paths to obtain a composite path that agrees with observations where they are available, agrees with the forecast well within the data gaps, and smoothly merges the forecast path to these observations at the edges of the data gaps. As part of these experiments, the experimenter reconciled such differences by adding or deleting points from the forecast path as needed without examining the known, gapless path for guidance. The role of the operator and the potential for an objective technique to assimilate the observations and a forecast are discussed more fully in section VIII.

In the proof-of-concept study, a persistence forecast assumes that the initial state for the forecast persists for the duration of the forecast. However, the initial states for these forecasts are composites of data and a forecast. Hence, an assumption that the initial state of the forecast is persistent is an ambiguous measure of how well the system with the forecast model compares to a system that has no such model.

An alternate definition of persistence is used here. A composite frontal path is a blend of data where it is available, an application of PATHFINDER across short data gaps, and an assumption that the previous week's path is persistent across long data gaps. PATHFINDER also blends the data and the persistent path. By applying this technique to all paths in a continuous series of experiments, the present capability is modeled.

The average absolute distance (error) between two curves over a specified longitudinal range is used as the quantitative measure in the evaluation of the forecasts. This distance is obtained using the program AXERR to divide the area between the curves by the path length. In particular, AXERR is used to determine the error between the known, gapless CIMREP path on one hand and the composite of data and the forecast. A comparison of this error to that between the gapless CIMREP path on one hand and a composite of data and an assumption of persistence on the other hand is used to evaluate the forecasts.

By using observed data gaps, by simulating the role of the operator, and by using the forecast in its intended role as a means of filling data gaps, these experiments are a realistic evaluation of the model in a quasi-operational mode. The objective of these experiments is to compare the total system, which is the DART model and PATHFINDER, to the present operational capability, which is PATHFINDER plus an assumption of persistence.

2. Example – 88/131

Figure 11 is a plot of the frontal path for 88/131 (11 May 1988); the IR data coverage was sufficient to map virtually the entire length of the path. After the eighteenth map (J.D. 354 of 1989) of data gaps is superimposed on this path, only the fraction of the path shown in Figure 12 can be seen. The previous week's (88/124) model initialization and 1- and 2-week forecasts are shown in Figure 13. The data available on 88/131 (Figure 12) and the 1-week forecast segment (initialized on 88/124 and valid on 88/131) that is used to fill the data gap are shown in Figure 14. PATHFINDER is used to merge the data and forecast segment, as well as to interpolate across short data gaps. The resulting

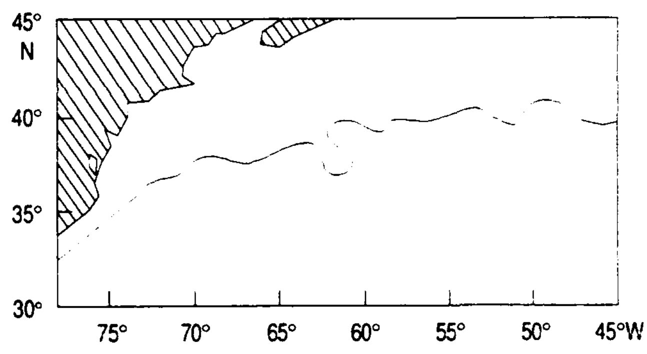

Figure 11. The path of the Gulf Stream frontal axis on Julian day 88/131. There were abundant data for analysis of the path in this case.

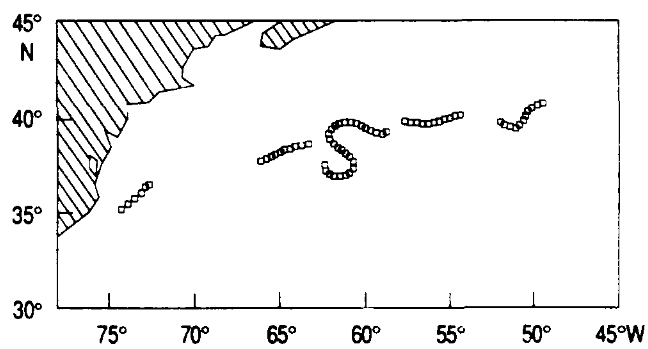

Figure 12. The observed positions of the Gulf Stream frontal axis on Julian day 88/131 after the 18th map of observed data gaps is superimposed on the path shown in Figure 11.

composite is shown in Figure 15. The observations available for the initialization of the previous forecast on Julian day 124 are shown in Figure 16.

Across the data gap, the composite of data and the forecast has a smaller error than does the composite of data and persistence (see Table 11). The errors are 12.2 km and 17.1 km, respectively. The improvement is 4.9 km, or 28.7% over the persistence error. A comparison of the gapless path to the two composites (Fig. 17), shows that the forecast composite has visually better agreement with the gapless path.

C. Ensemble of Experiments

1. Summary

Table 11 presents the results of an ensemble of data assimilation experiments. This table shows the error between the gapless observed path and the forecast composite, as well as that between the observed path and the persistence composite. Also shown is the difference between the two errors, as well as the percent difference. The forecast composite has a smaller error than does the persistence composite; i.e., the former is a better composite in 10 of the 14 cases, or 71% of the cases. In two of the cases in which the persistence composite is better (87/104, gap map 18, and 88/138, gap map 20), the difference is less than 2 km. In the other two cases of the persistence composite being better than the forecast composite (89/158, map 23, and 89/165, map 24), the result can be attributed to a problem with the forecast initialization process.

A scatter plot of the forecast and persistence errors is shown in Figure 18. The diagonal line across the plot indicates where the forecast and persistence errors are equal. Most of the points lie close to or below the line. The two points that lie the farthest above the line are the two cases that have the forecast initialization problem.

A histogram of the forecast and persistence errors is presented in Figure 19. It shows that while the persistence errors are 30 km or larger in five cases, the forecast errors are this large in only three

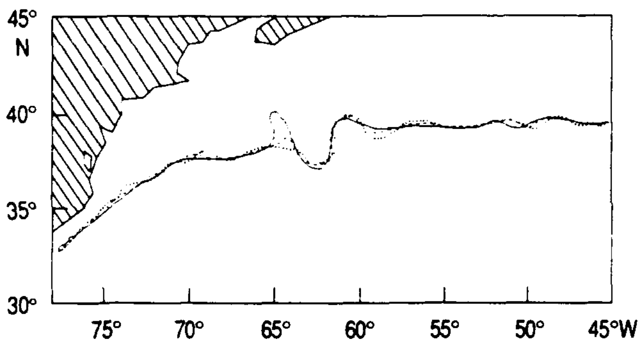

Figure 13. The initial state (solid line), 1-week forecast (dashed line) and 2-week forecast (dotted line) of the frontal axis for a forecast initialized on Julian day 124 of 1988. The initial state is obtained by superimposing the 17th map of observed data gaps on the CIMREP frontal path data Julian day 124 of 1988 and using PATHFINDER to interpolate across the data gaps.

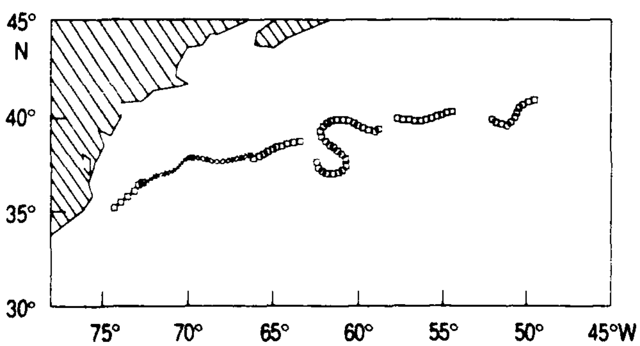

Figure 14. The frontal path observations available on Julian day 131 of 1988 (□, the same as Fig. 12) and the 1-week forecast path valid on Julian day 131 of 1988 (x) for the long data gap at 66°–73°W.

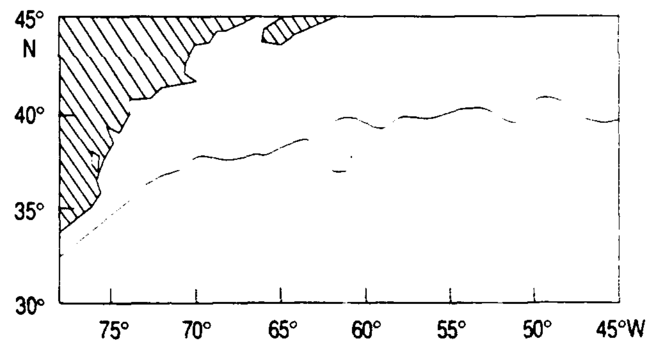

Figure 15. The frontal path obtained by using PATHFINDER to fit a path to the data shown in Figure 14. All of the observations and forecast positions are assumed to be new (no aging).

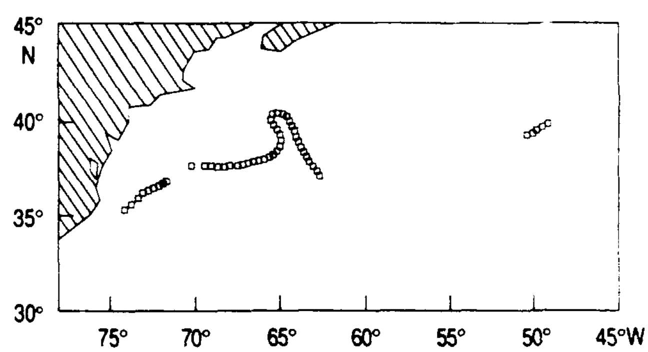

Figure 16. The observations available for the initialization of the model forecast on 88/124 (Fig. 13).

Table 11. Average absolute offset error for composites of data and a forecast, as well as composites of data and an assumption of persistence for an ensemble of cases. The errors are relative to the gapless observed path and are calculated across superimposed data gaps.

DATE (YR/J. D.)	DOMAIN (LONG. W)	GAP MAP	PERSISTENCE ERROR (km)	FORECAST ERROR (km)	(P-F) (km)	(P-F)/P %
87/104	72.8–66.1	18	21.1	22.9	−1.8	−8.5
87/111	58.0–51.0	19	22.8	14.4	8.4	36.8
87/133	66.1–72.8	18	56.5	42.2	14.3	25.3
88/131	66.1–73.0	18	17.1	12.2	4.9	28.7
88/138	51.0–58.2	19	57.9	56.6	1.3	2.2
88/131	51.0–58.2	19	32.4	28.9	3.5	10.8
88/138	64.9–75.0	20	16.0	16.5	−0.5	−3.1
88/138	51.0–61.5	20	37.2	30.5	6.7	18.0
89/158	66.1–73.0	18	23.5	22.9	0.6	2.6
89/165	51.0–58.2	19	22.7	22.6	0.1	0.4
89/172	64.9–75.0	20	31.6	25.1	6.5	20.6
89/172	51.0–61.5	20	16.6	14.6	2.0	12.0
89/158	54.7–57.0	23	12.1	18.9	−6.8	−56.2
89/165	54.8–57.9	24	19.3	26.2	−6.9	−35.8
		Averages	27.6	25.3	2.3	8.3

cases. The forecast composites also have very small errors (10–15 km) in three cases, while the persistence errors are this small in only one case.

The averaged results are also presented in Table 11. The ensemble of persistence cases has an average error of 27.6 km. The average error for the forecast cases is 25.3 km, which is an improvement of 2.3 km, or 8.3% over the average persistence error. The improvement in the error of the forecast composites over that of the persistence composites can be used to judge the skill of the forecast system over persistence (Fig. 20). In six cases, the improvement is greater then 3 km (approximately 10% of the persistence error), in six cases it is

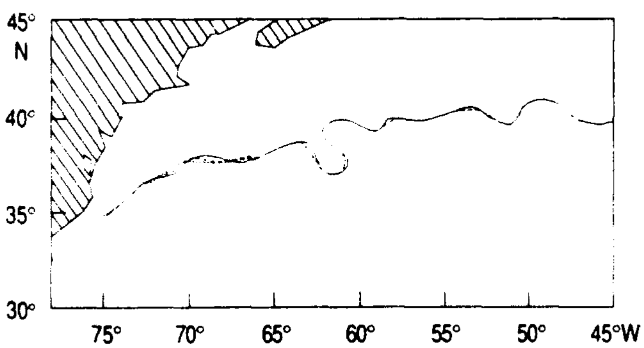

Figure 17. A comparison of the observed, gapless CIMREP frontal path for 88/131 (solid line), the composite of available observations and a 1-week forecast (dashed line) and a composite of the available observations and an assumption of persistence (dotted line).

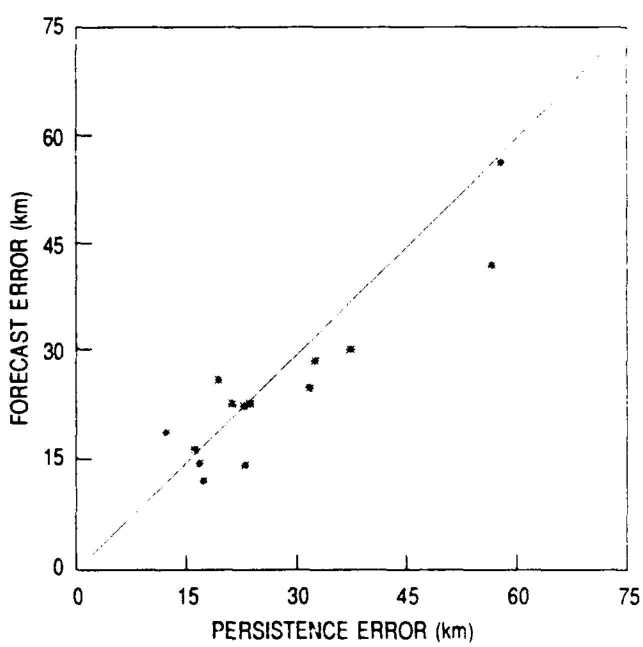

Figure 18. A scatter plot of the error of the composite of data and a 1-week forecast to the error of a composite of data and persistence. Both errors are relative to the gapless, CIMREP path.

between plus and minus 3 km, and in just two cases, it is worse than a degradation of 3 km. If a percent change in the forecast error over the persistence error of 10% or larger is arbitrarily judged to be significant, then the forecast composites are significantly better than the persistence composites in one-half of the cases, are comparable in about one-third of the cases, and are significantly worse in the remaining cases (Fig. 21).

2. Example – 89/158–89/165

The use of this forecast system requires that the path of the Gulf Stream north wall and that of the frontal axis be inferred from each other. The reason for this inference is that the forecast model determines the path of the axis, which is readily merged with the CIMREP axis paths, but the feature model used in the model initialization process keys off the north wall. In the implementation used for the proof-of-concept study, as well as for these experiments, the north wall is assumed to lie 40 km to the left of the axis while looking downstream. The feature model in OTIS is used to construct a three-dimensional thermal analysis from the path of the north wall. The surface topography of this thermal field is then estimated, and the axis of the front in that topography field determined. If the north wall is inferred from an axis path and the axis then inferred from that north wall, then the inferred axis does not necessarily agree with the original axis wherever the path has curvature. This shifting procedure is specific to the experiments using the CIMREP paths, which are axis paths. The operational boguses are north wall paths. Experiments using operational boguses have a different, but related problem of inferring the north wall path from the model-generated axis path so that it can be merged with the north wall observations.

The following example demonstrates the consequences of this lack of agreement. Figure 22 shows

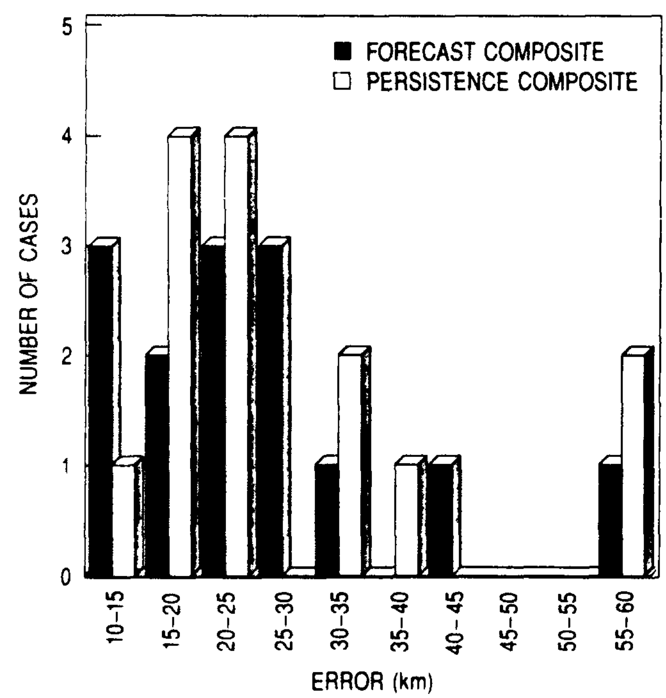

Figure 19. A histogram of the distribution of error within the ensemble of composites of data and a 1-week forecast as well as the corresponding ensemble of composites of data and persistence.

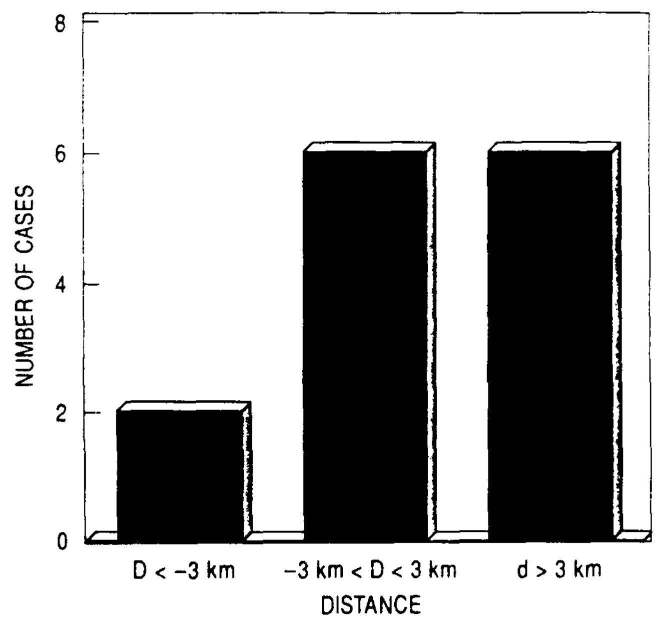

Figure 20. A histogram of the difference in error (persistence composite error minus forecast composite error).

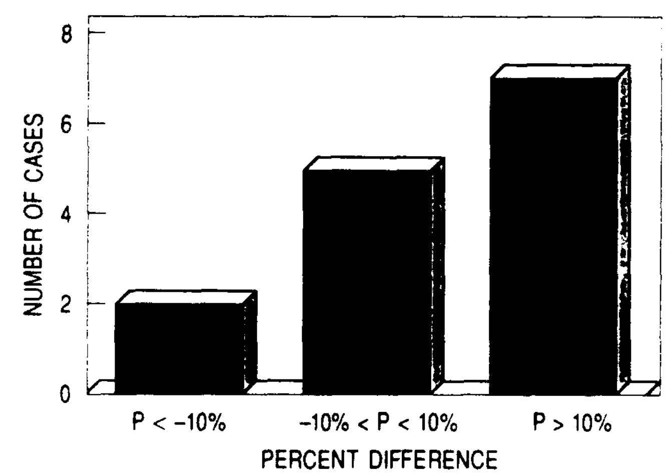

Figure 21. A histogram of the percent difference in error ((persistence composite error minus forecast composite error)/persistence composite error).

the observed, gapless frontal paths on Julian days 151 (31 May), 158 (7 June) and 165 (14 June) of 1989. A data gap is superimposed on Julian days 158 (54.7°–57.0°W) and 165 (54.8°–57.9°W). These two gaps are nearly identical and cover much of a meander at 55°–56°W. The observations show that this meander maintains its initial amplitude and perhaps broadens a little. In addition, the eastern wall of the meander is observed to deepen southwestward. Figure 23 shows the composites of data (outside of the gaps) and the previous 1-week forecast (across the gaps) for Julian days 158 and 165. The composite meander maintains its amplitude and deepens southwestward. However, the amplitude is substantially smaller than observed. As a result the statistical analysis presented above shows that persistence is a better description of the meander than the composites are.

The reason for the small amplitude is the problem with shifting between the axis and north wall. Figure 24 shows the observed frontal axis on Julian day 151, as well as the axis of the model initialization generated from that path. Similarly, Figure 25 shows the 1-week forecast valid on Julian day 158 and the axis of the model initialization of that day. (Note that the Julian day 158 initialization uses a composite of data outside of the gap and the previous 1-week forecast across the gap, which covers the meander.) In both cases, going from an observed or forecast axis to the north wall and then to the axis of the feature-modeled field results in a reduction of the amplitude of the meander.

3. Statistical Significance

The same statistical analysis used to test the significance of the experiments with operational boguses (section VIIB), is applied to these experiments. The results are shown in Table 12. In this case, the distribution is sufficiently Gaussian to use the t-test. The mean, 2.31 km, is larger than the standard error in the mean, which is 1.53 km. The significance level is 0.156, which means that we can be 84% confident that the forecast composite error is an improvement over that of the persistence composite.

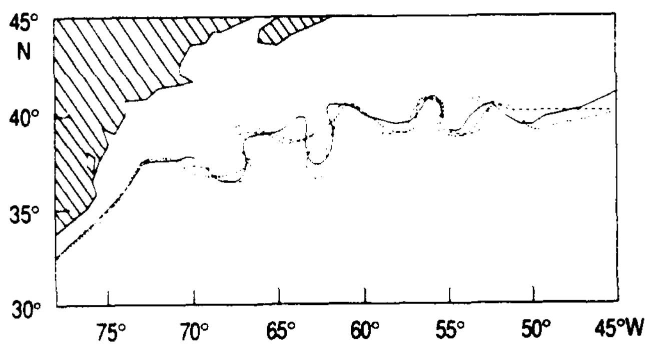

Figure 22. A plot of the gapless, CIMREP frontal paths for 89/151 (solid line), 89/158 (dashed line) and 89/165 (dotted line).

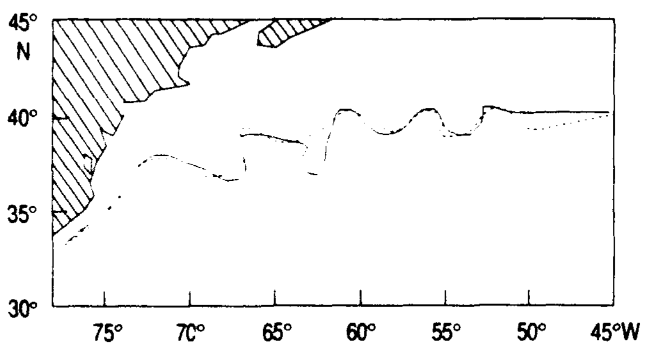

Figure 23. A plot of the composites of data and a 1-week forecast for 89/158 (solid line) and 89/165 (dashed line). The data gaps are at 54.7–57.0 on 89/158 and 54.8–57.9 on 89/165.

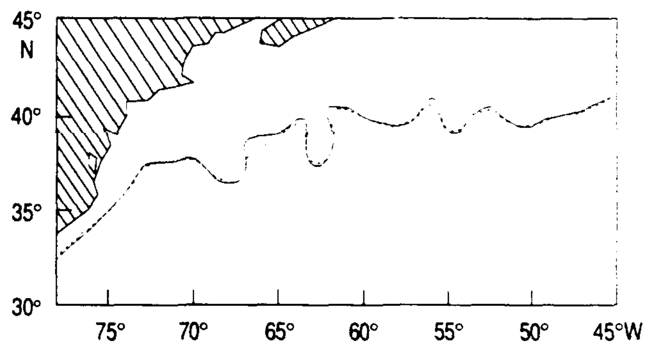

Figure 24. A comparison of the observed, gapless frontal path on 89/151 (solid line) and the axis of the model initialization generated for the same date (dashed line). The two paths disagree near 63.3–68.9W because of a data gap. However, there is no data gap at 54–58W at this time.

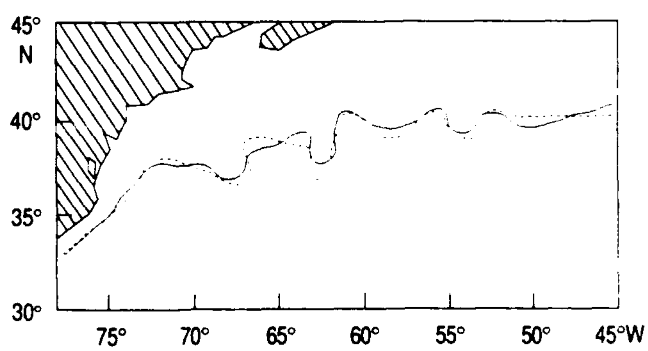

Figure 25. A comparison of the 1-week forecast valid on 89/158 (solid line) and axis of the frontal path in the forecast initialization on that date (dashed line).

Table 12. Summary of statistical analysis of the difference in error between the persistence composite and the forecast composite (persistence minus forecast).

	Mean	= 2/31 km
	Std. Dev.	= 5.73 km
	Std. Err.	= 1.53 km
Shapiro-Wilk test:		
	W(Normal)	= 0.969
	Prob < W	= 0.817
T-test:		
	T(Mean = 0)	= 1.51
	Significance	= 0.1561
	Confidence	= 84%

Also presented are the results of the Shapiro-Wilk test, which shows the distribution to be sufficiently Gaussian, and the results of the T-test. These results indicate that the forecast composite has a mean improvement in the error of 2.3 km over that of the persistence composite. Using a T-test, we have a confidence of 84% that using the forecast rather than persistence improves the frontal nowcast.

VII. Evolution of Initialization Uncertainties

A. Parallel Experiments

There are numerous sources of uncertainty in the initial state of the forecast. Digitization of a frontal path from abundant IR data by even a skillful operator has uncertainties. An ensemble of digitizations would form an envelope with an amplitude of perhaps 15–20 km about the mean path of the ensemble. If there is a gap in the path data, then the operator must either make a subjective interpolation across the gap or use some interpolation software. Both alternatives introduce uncertainties.

How do these uncertainties evolve with time? This question is investigated by comparing pairs of parallel forecasts. A gapless CIMREP path is used to initialize both forecasts of a given pair. However, for one of the forecasts, a section of the path is deleted and the PATHFINDER system is used to interpolate across the resulting gap. No data, such as a previous forecast, are used to aid the interpolation.

Figure 26 is an example of a model initialization from the CIMREP path dated 8 April 1987 (87/097) as well as the 1- and 2-week forecasts (the standard case). Figure 27 is a parallel initialization and forecast (case a) but with a section (67°–70°W) of the initial standard path deleted, and PATHFINDER used to interpolate across the gap. The initializations and forecasts for the two cases are shown side-by-side in Figures 28–30. In this case, the initial differences are on the western side of a trough in the standard case. Figure 28 shows that the data at the western edge of the gap is propagated downstream about one correlation scale length and then smoothly joined to the data on the eastern edge of the gap. The result is a trough in the path that is narrower than and out of phase with the standard case. The time evolution of the paths shows that at 2 weeks the trough of case a is about 90° out of phase with the trough of the standard case.

The latitudinal offset of the paths is shown in kilometers in Figure 31. The initial difference

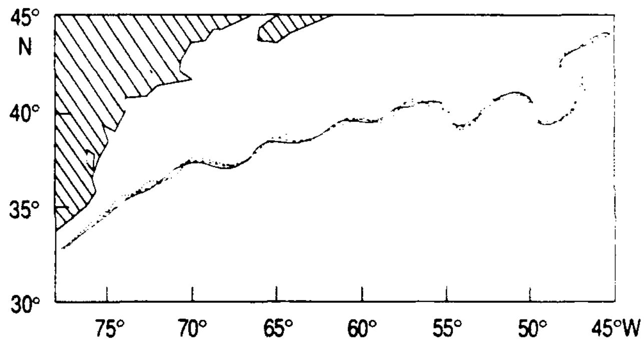

Figure 26. A forecast model initialization (solid line), a 1-week forecast (dashed line) and a 2-week forecast (dotted line). The initialization is dated 8 April 1987. This is the standard case for the following comparisons.

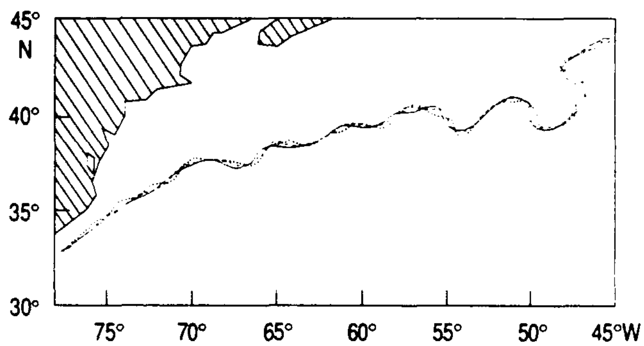

Figure 27. A forecast model initialization (solid line), a 1-week forecast (dashed line) and a 2-week forecast (dotted line). The initialization is dated 8 April 1987. In this perturbed case, the standard initial state (Figure 26) has been modified by deleting a section of the front near 67–70W and using PATHFINDER to interpolate across the resulting gap.

between the paths can be seen translating downstream (to the right) with an amplitude that slowly decreases with time. Figure 32 is a plot of the same curves but in a frame of reference that translates downstream at a speed of 38 cm s^{-1}. The initial difference is virtually at rest in this frame of reference. In addition, waves of short wavelength can be seen dispersing about the initial difference. The initial difference has a small negative extremum on its leading edge near 66°W. This extremum is substantially larger at later times and is propagating downstream. In addition, a wavelike structure can be seen developing behind the leading edge and is especially prominent at 2 weeks. A similar evolution can be seen for the small positive extremum in the initial difference near 62.3°W. In both cases, the speed of the these short waves is about 9 cm s^{-1} (in the translating reference frame) plus 38 cm s^{-1} (the speed of the reference frame), or about 47 cm s^{-1}. Finally, the leading edge of

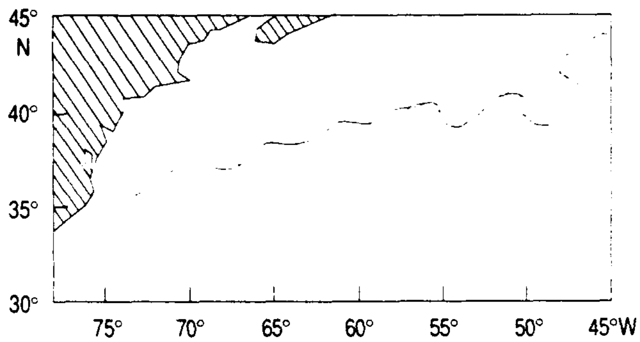

Figure 28. A comparison of the initial states of the standard case (solid line) and case a (dashed line). See Figures 26 and 27.

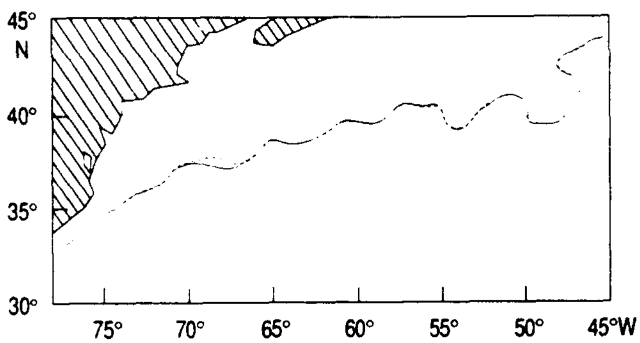

Figure 29. A comparison of the 1-week forecasts of the standard case (solid line) and case a (dashed line). See Figures 26 and 27.

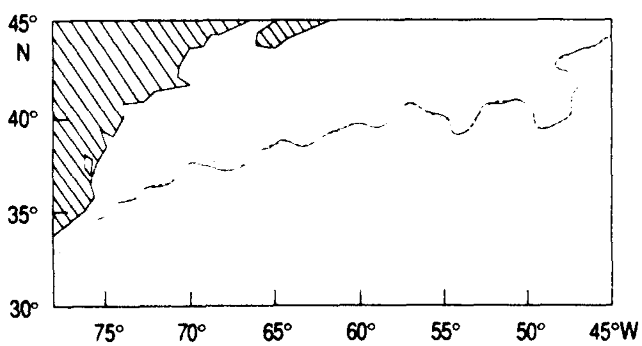

Figure 30. A comparison of the 2-week forecasts of the standard case (solid line) and case a (dashed line). See Figures 26 and 27.

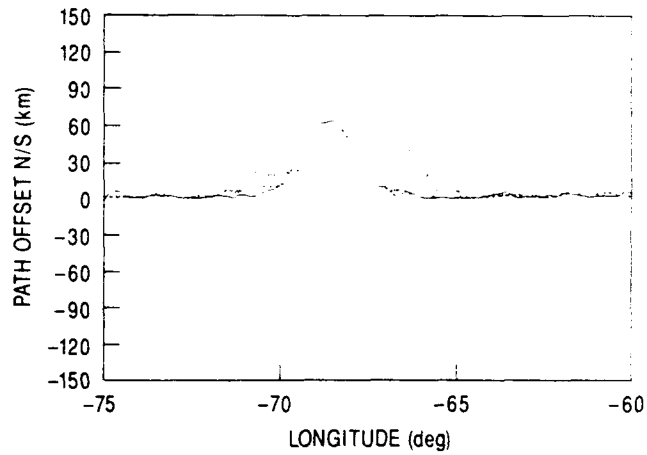

Figure 31. A comparison of the latitudinal offset in kilometers between case a and the standard case (case a – standard) at the time of initialization (solid line), at 1 week (dashed line) and at 2 weeks (dotted line).

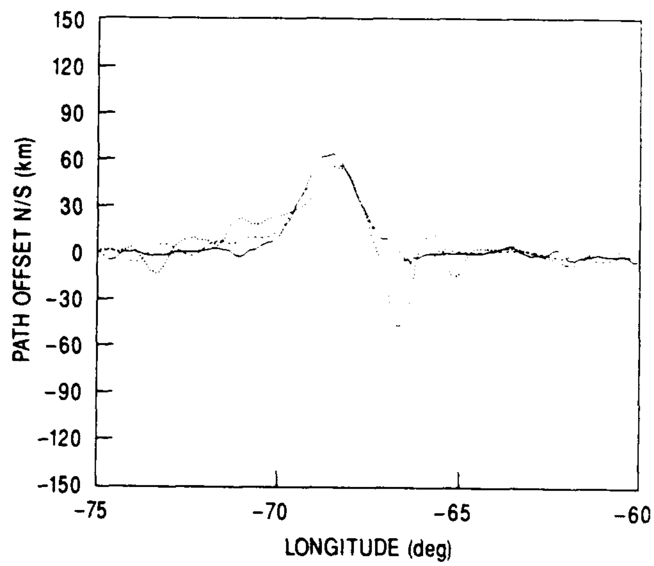

Figure 32. A comparison of the latitudinal offset in kilometers between case a and the standard case (case a – standard) at the time of initialization (solid line), at 1 week (dashed line) and at 2 weeks (dotted line) in a frame of reference that is moving eastward (right) at a speed of 38 cm s^{-1}.

initial difference between the two paths develops a sinc$(x) = \sin(x)/x$-like structure.

Cases using gaps at different locations and of different lengths have also been made. Figure 33 presents the initializations of case b, in which the difference lies on the eastern side of a trough in the standard case. Figure 34 shows the differences between the standard and case b. Again, the evolution resembles a difference field that is translating downstream at a speed of approximately 38 cm s^{-1} with short dispersive waves developing both upstream and downstream of the initial difference. Note the evolution of the leading edge of the initial differences, as well as the extrema in the initial differences near 72.2°W and 62.2°W. Another example, case c, is presented in Figures 35 and 36. Figure 36, which is a plot of the differences in a frame of reference that is translating downstream at 38 cm s^{-1}, suggests that the peak differences translate downstream at approximately 38 cm s^{-1}, but that the main body of the differences is translating at a somewhat slower speed. The sinc(x) structure of the leading edge and the dispersive waves are clearly evident in this case.

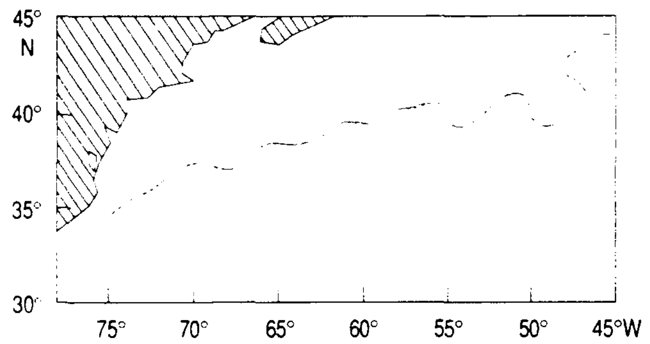

Figure 33. A comparison of the initial states of the standard case (solid line) and case b (dashed line). In case b, the standard initial state was deleted between about 66 and 68W, and PATHFINDER was used to interpolate across the resulting gap.

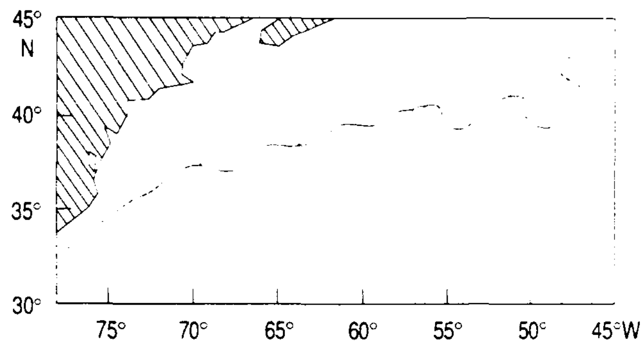

Figure 35. A comparison of the initial states of the standard case (solid line) and case c (dashed line). In case c, the standard initial state was deleted between about 64 and 70W, and PATHFINDER was used to interpolate across the resulting gap.

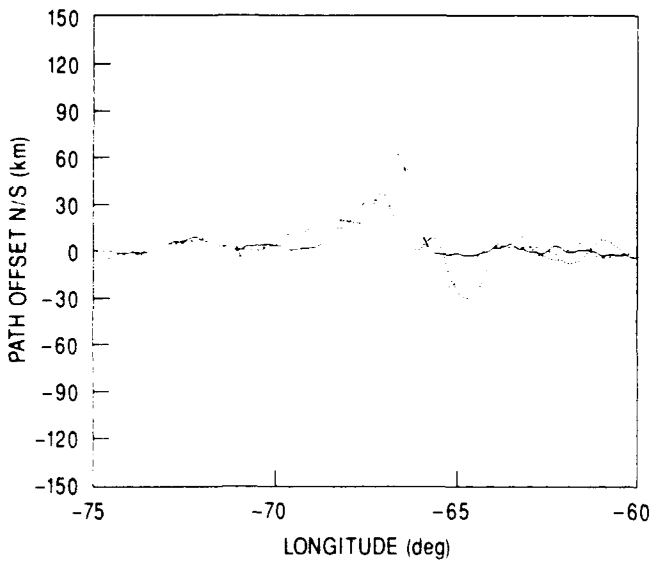

Figure 34. A comparison of the latitudinal offset in kilometers between case b and the standard case (case b – standard) at the time of initialization (solid line), at 1 week (dashed line) and at 2 weeks (dotted line) in a frame of reference that is moving eastward (left) at a speed of 38 cm s^{-1}.

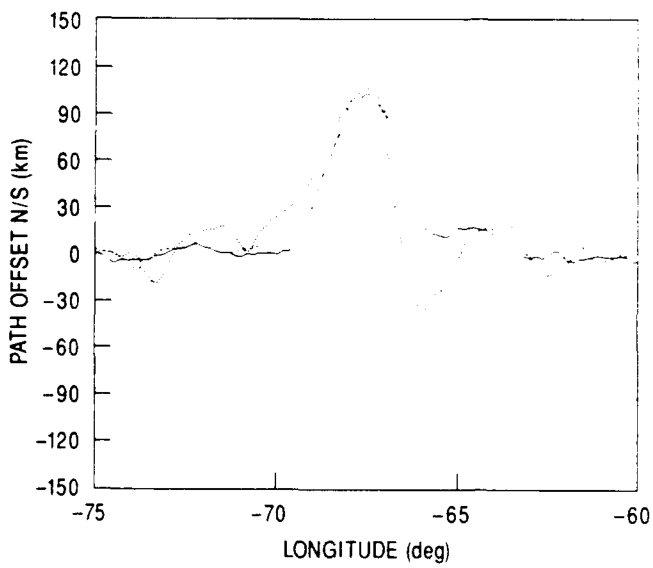

Figure 36. A comparison of the latitudinal offset in kilometers between case c and the standard case (case c – standard) at the time of initialization (solid line), at 1 week (dashed line) and at 2 weeks (dotted line) in a frame of reference that is moving eastward (left) at a speed of 38 cm s^{-1}.

B. Monte Carlo Estimates of Forecast Confidence

Fox et al. (1990) describe a Monte Carlo approach to estimating the confidence in the forecast, which is summarized here because of its relevance to the operational use of the model. Given an initial Gulf Stream frontal path, an ensemble of initial states within an envelope of uncertainty can be obtained by applying a rubber sheet error model to the path. This error model distorts the path by shifting the path in a random but correlated manner. If a forecast is made for each initial state in this ensemble, then the relative size of the forecast envelope at frontal points could give an estimate of the sensitivity of the forecast path to uncertainties in the initial state. This approach could identify frontal segments that require special care in their analysis.

VIII. Role of the Operator

Although the forecast system can be run by having the operator execute three sets of commands, called UNIX scripts, plus graphics as desired, the assimilation of the forecast and observations into a composite analysis of the frontal path could require an operator to interact with the assimilation process. In particular, the issue is how to reconcile the forecast and observations at the edge of the data gaps. The PATHFINDER system has not been well tested for this task. For the experiments presented in section VI, an engineering-like approach was used to combine PATHFINDER and the available operator into a workable technique that did not require extensive testing. In this section, some examples of this approach are presented as performed by the experimenter. These examples identify cases of mismatched observations and forecasts that any technique will need to handle.

The Technical Validation Panel discussed the proper way to blend the forecast and the observations into a composite path. They noted that in atmospheric modeling, the forecast is used as a first guess and the observations are blended into it via optimum interpolation. The suggestion was made that given the proper relative weights, PATHFINDER could indeed blend frontal path observations into the forecast path without generating unphysical paths. A first estimate of the relative weights could be made by computing the ratio of the uncertainty in the observations to the uncertainty in the forecast. The uncertainty in the analysis of IR images, about 15 km, could be assigned to new, 1-day-old observations; aged data would have a larger uncertainty. The mean errors of the forecasts initialized and verified with the operational boguses (Table 3) could be used as a first estimate of the forecast uncertainty. The resulting relative weights are presented in Table 13. Several issues need to be examined: how well can PATHFINDER blend the forecast and observations, what are the proper relative weights, and what are the proper spatial and temporal correlation scales.

A. Example – 88/138

The observed frontal path and the model forecast do not necessarily agree at the edge of a data gap. Figure 37 shows the observed gapless evolution of the frontal path from 88/124 (4 May 1988) to 89/138 (18 May 1988). Note in particular the major changes that are observed in the path between 60° and 65°W. With the 20th map of data gaps superimposed, only a small fraction of the path is seen on 88/138 (see Fig. 38). The 88/131 initialization of the model and the 1-week forecast (valid 88/138) are also shown on Figure 38. The model missed the pinch-off of a cold-core ring and the subsequent northward shift of the path. As a consequence, the

Table 13. First estimate of relative weights to be assigned to a forecast and new (1-day old) observations in an optimum interpolation-based technique to combine a forecast path and path observations into a composite path.

	WESTERN AREA (73°–66°W)	CENTRAL AREA (66°–59°W)	EASTERN AREA (59°–53°W)
Observations	1.0	1.0	1.0
Forecast	0.85	0.39	0.38

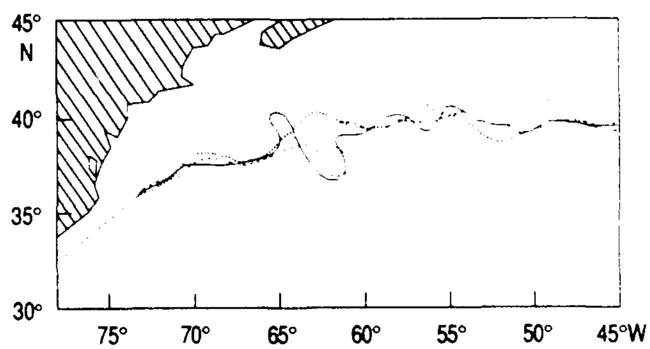

Figure 37. A plot of the gapless CIMREP frontal paths for 88/124 (solid line), 88/131 (dashed line) and 88/138 (dotted line).

observed and forecast paths are offset from one another at the edge of the data gap at 65°W. To smoothly blend the two paths, the forecast path is deleted between 65°W and 66°W. PATHFINDER is then used to smoothly interpolate between the forecast and observed paths. Figure 39 shows the edited forecast segment, the observed path, and the composite path. The composite path agrees with the observations where they are available, agrees with the forecast away from the data, and smoothly merges the observations with the forecast.

B. Example – 87/111

Sometimes, a composite meander must be constructed. Figure 40 is an example. After the 19th data-gap map is superimposed on the complete CIMREP path for 87/111 (21 April 1987), only the observations shown by the squares are available; there are no observations east of 58°W. The 1-week forecast valid on 87/111 for the area east of 58°W is shown by the x's. The paths do not agree well at 58°W. The path fitted to these data by PATHFINDER is shown in Figure 41; it is not realistic. The forecast positions between 57° and 58°W are deleted (Fig. 42) in the hope that PATHFINDER will smoothly connect the observations to the remaining forecast positions. However, the fitted path (Fig. 43) extends farther north than would be expected. By inserting an additional point in the forecast path (shown by the triangle in Fig. 44), a fitted path is obtained that is a subjective best fit of the data and the observations. This path fits the observed amplitude and the western wall of the meander, as well as the forecasted position of the eastern wall. A comparison of the composite meander (Fig. 45) to the gapless verification path (also shown on Fig. 45) shows excellent agreement.

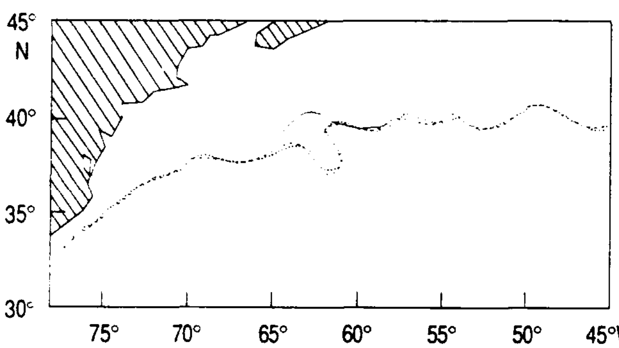

Figure 38. A plot of the available frontal path data after the 20th map of data gaps has been superimposed on the CIMREP path for 88/138 (solid line), the forecast initialization dated 88/131 (dashed line) and the 1-week forecast valid 88/138 (dotted line).

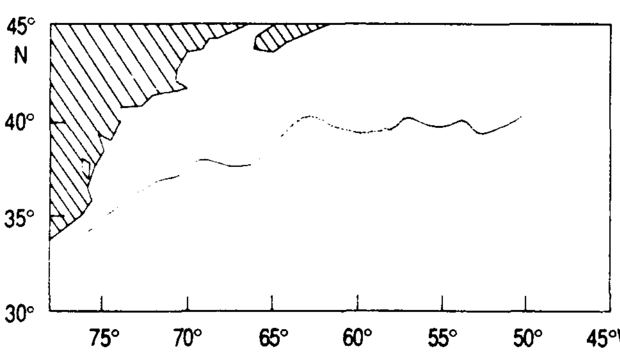

Figure 39. A plot of the edited forecast segments that are used for filling the data gaps (solid lines), the available observations of the frontal path (dashed line) and the path fitted to the forecast and observations by using PATHFINDER (dotted line).

C. Example – 89/172

Not all instances of operator involvement are successful. Figure 46 shows the available observations and the forecast path segments for 89/172

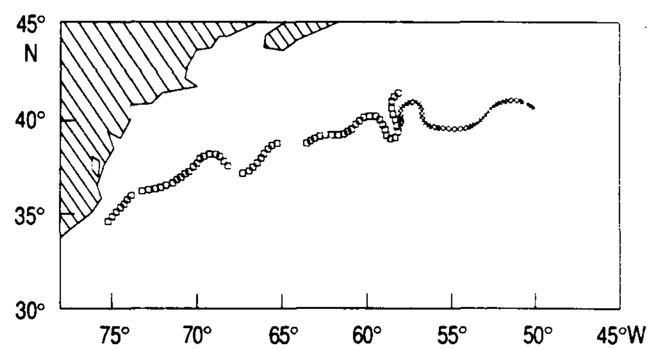

Figure 40. A plot of the available frontal path observations (□) and the 1-week forecast segment used to interpolate across the data gap east of 58W (x). These data and forecast positions are valid 87/111.

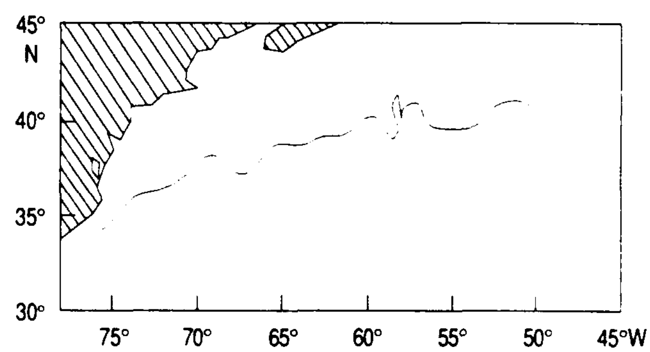

Figure 41. The frontal path obtained by using PATHFINDER to fit the data and forecast in Figure 40.

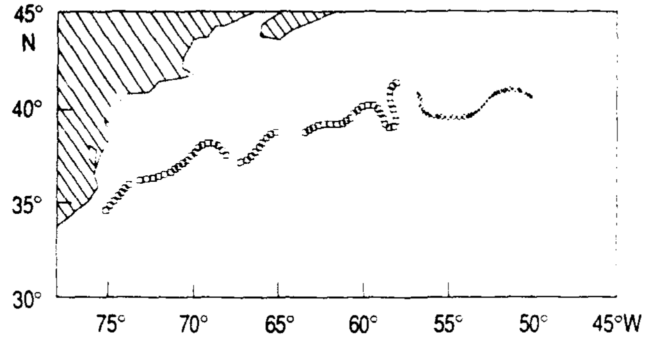

Figure 42. The same as Figure 40 but with the edited forecast segment. The western edge of the forecast segment has been edited eastward.

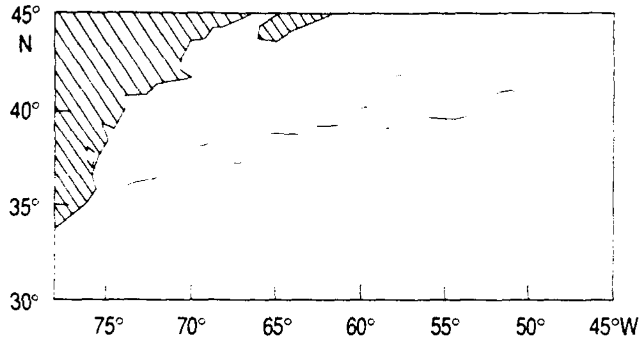

Figure 43. The frontal path obtained by using PATHFINDER to fit the data and forecast in Figure 42.

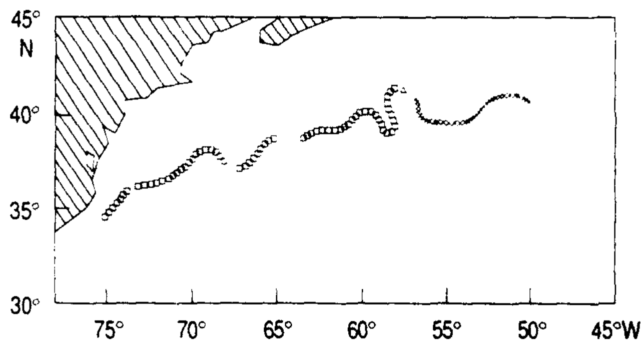

Figure 44. The same as Figure 42 but with a "bogus" point (the triangle) inserted as additional data.

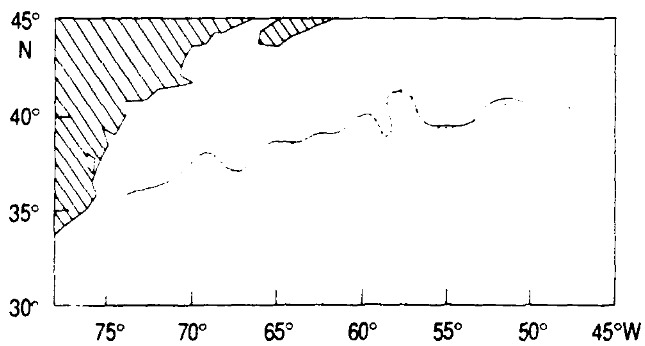

Figure 45. A comparison of the composite of available observations and a 1-week forecast (solid line) obtained by applying PATHFINDER to the data in Figure 44 to the gapless, CIMREP verification path (dashed line).

(21 June 1989). There is a mismatch between the data and the forecast near 65°W. To reconcile this mismatch, the eastern end of the forecast segment is deleted (see Fig. 47) and PATHFINDER is used to merge the remaining forecast path to the observations. The resulting composite path (also shown in Fig. 47) has a broad meander and the pointed corners of both the data and the forecast. In contrast, the gapless paths of 89/165 and 89/172 (Fig. 48)

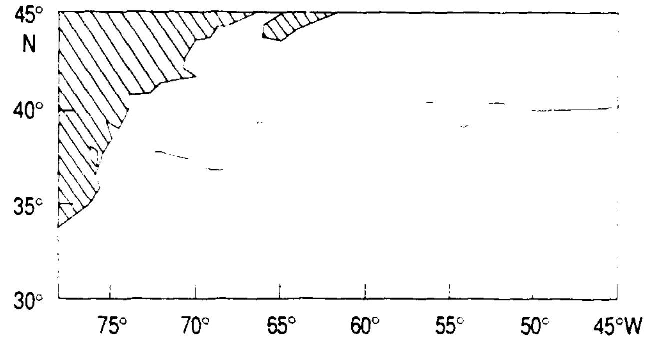

Figure 46. A plot of the 1-week forecast segments (solid lines) and the available path observations (dashed line) for 89/172.

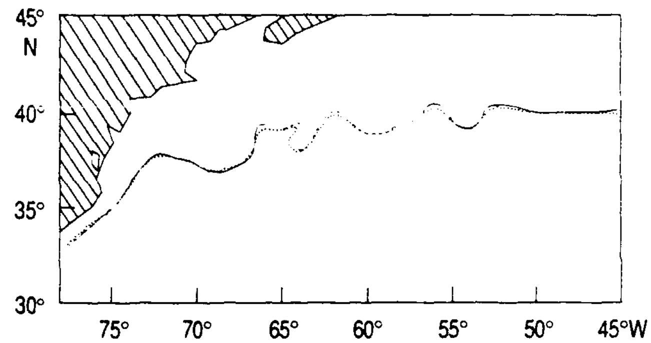

Figure 47. The same as Figure 46 but with the edited forecast segments. The ends of the forecast segments that are adjacent to the observations have been edited away from the observations. PATHFINDER has been used to fit a curve to these data and forecast segments (dotted line).

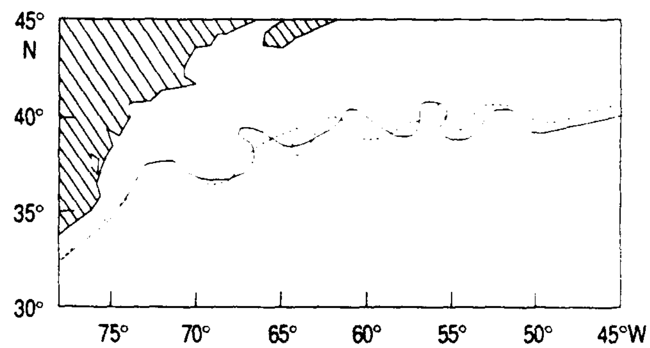

Figure 48. A plot of the gapless, CIMREP frontal paths for 89/165 (solid line) and 89/172 (dashed line).

show only one such corner on 89/172. Specifically, Figure 48 shows that the axis of the sharp meander near 65°W rotates clockwise about 90° during the week. The 1- and 2-week forecasts initialized on 89/165 (Fig. 49) show this sort of behavior but at a slower rotation rate. Hence, the meander appears in different locations in the forecast and observations. The edge of the data gap lies between them so that both meanders appear in the composite. In hindsight, the operator should have deleted the entire meander from the forecast path and used PATHFINDER to smoothly merge the remaining forecast path with the observations. With experience, an operator might be able to compare the available observations to the model forecast and, realizing that the model is perhaps a little slow in evolving features, make such a judgment.

D. Example – 87/133

In some cases, the operator might choose to interpolate across data gaps using a 2-week forecast rather than a 1-week forecast because the forecast model is slow in its modeling of events. Figure 50 shows the gapless observed path for 87/133 (13 May 1987), as well as the forecast and persistence composites. The deep meander that was seen near 67°–68°W on 87/126 (see the persistence path) has pinched off by 87/133, as shown by the gapless observations. The 1-week model forecast that is valid on 87/133 has not pinched off this ring. As a result, the forecast composite strongly resembles persistence. However, if a composite (see Fig. 51) is made using the 2-week forecast that is valid 87/140, then the composite agrees exceptionally well with the observations. The error between the observed path and the composites is dramatically reduced in the case using the 2-week forecast (Table 14). Operationally, an operator could examine the forecast for events that occur between the 1- and 2-week forecasts. If he or she finds such an event and the event lies with a data gap, then the 2-week forecast, rather than the 1-week forecast, could be used in making the composite.

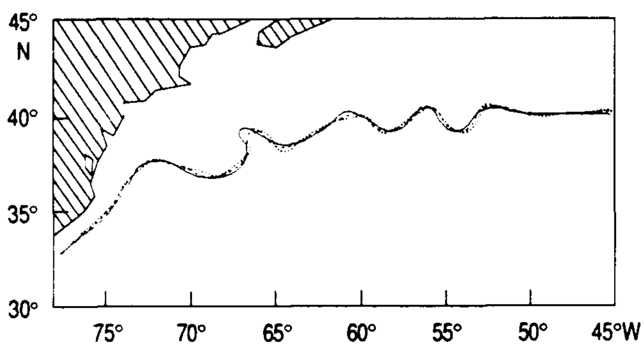

Figure 49. A plot of the axis inferred from the forecast initialization on 89/165 (solid line) and the 1-week (dashed line) and 2-week (dotted line) forecasts.

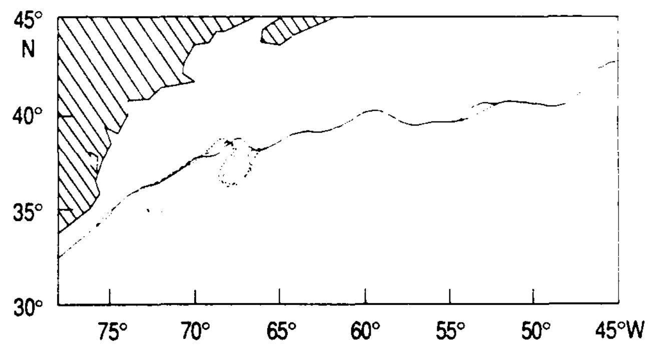

Figure 50. A plot of the gapless, CIMREP path for 87/133 (solid line) as well as the forecast (dashed line) and persistence (dotted line) composites.

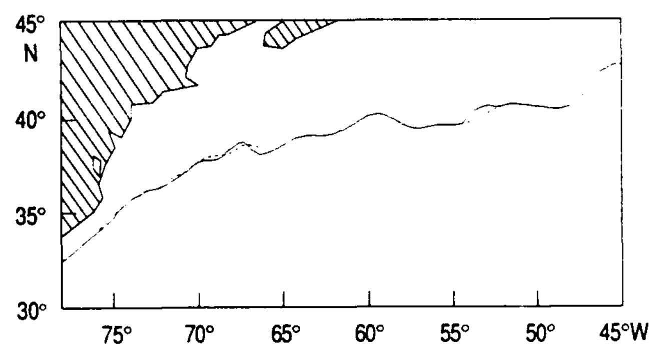

Figure 51. A plot of the gapless, CIMREP path for 87/133 (solid line) and a composite of the available observations and the 2-week forecast valid on 87/140 (dashed line).

Table 14. Comparison of error for a pair of data assimilation experiments.

FORECAST AS DATA	DATE VALID	PERSISTENCE ERROR (km)	FORECAST ERROR (km)	(P-F)/P (km)	(P-F)/P (%)
1-Week Experiment	13 May	56.5	42.2	14.3	25.3
2-Week Experiment	20 May	56.5	12.6	43.9	77.7

IX. Summary and Conclusions

A. Summary

The purpose of this work is to transition the DART Gulf Stream forecast system to operational U.S. Navy use. Recall that the proof-of-concept study demonstrated that the system makes 1- and 2-week forecasts of the Gulf Stream path that are statistically better than persistence when there are abundant data for initialization and verification. The results of an evaluation of the model in a more operationally realistic mode are presented here. Two sets of experiments designed to evaluate the forecast skill of the system in a quasi-operational environment are the crux of the report. In one set of experiments, operational front and ring maps are used to initialize and verify 1- and 2-week forecast experiments. In the other set of experiments, a simple data assimilation system is set up by using a 1-week forecast to interpolate across long data gaps. The gappy paths are constructed by superimposing observed data gaps on complete frontal paths. This data assimilation system is compared to the existing operational capability. This report also considers the evolution of uncertainties in the initial state. Finally, the potential role that an operator would have while using this system as an aid for constructing a front and ring map is described.

The DART Transition Plan, (see appendix) listed six thrusts for evaluating the DART system.

- forecast sensitivity to data gaps
- real-time forecast sensitivity studies
- the value added by the forecast system
- new tests using abundant data
- sensitivity of the model initialization to the version of OTIS
- other measures of forecast performance.

B. Conclusions

The DART Gulf Stream forecast system can make 1- and 2-week forecasts of the Gulf Stream frontal path using the operational boguses for initialization. These boguses are better than an assumption of persistence. The forecast errors are typically 10% and 15% better than persistence errors at 7 and 14 days, respectively. The results of a statistical analysis indicate that the forecast has a smaller error than persistence: the statistical significance ranges from 92% at 1 week in the eastern domain to over 99% at 2 weeks in the western and center domains. This work addresses thrust 2 of the transition plan.

An ensemble of data assimilation experiments shows that frontal path nowcasts that are composites of data and a forecast are typically comparable to or better than composites of data and persistence. Results show that the errors of the forecast composites are on average 2.3 km, or 8.3%, better than the present operational capability. Using a t-test, we have a confidence of 84% that using the forecast rather than persistence improves the frontal nowcast. This work addresses thrust 3 of the plan.

Pairs of experiments show that the difference in initial states is propagated downstream at approximately 38 cm s^{-1}. Dispersive short waves are evident. The leading edge of the difference develops a $\mathrm{sinc}(x)$-like structure. This work addresses thrust 1 of the plan.

The inference of the frontal axis and the north wall from each other, which is part of the total system but not the forecast module itself, is a technical problem that degraded several forecast composites. It should be noted that the frontal axis is determined from the pressure field and that the north wall is observed by examining thermal gradients at the surface; hence, these paths are quite distinct entities.

The assimilation of the forecast and the observations into a composite nowcast of the frontal path often requires that a mismatch between the forecast and the observations at the edge of a data gap be reconciled. In the work reported here, we applied an engineering approach that uses available resources and experience rather than a new development effort. In particular, an operator reconciles the mismatch by editing the forecast. Cases of special concern are those in which the forecast has the correct character but is slow in evolving features. A trained operator can use this approach to successfully reconcile these mismatches.

The research-grade boguses for new tests using abundant data were not available in time for these tests to be performed (thrust 4 of the plan). Some experiments were made using boguses from periods of abundant IR data in 1989, and the results were comparable to the proof-of-concept results. The transition work was directed to use OTIS 2.1, which was used in the proof-of-concept study; thrust 5, therefore, was not addressed. The latitudinal offset between parallel forecasts should be considered in the study of the evolution of the initial uncertainty. Otherwise, the average absolute distance (error) will be the quantitative measure.

X. Recommendations

Couple the DART system and the PATHFINDER system. Pursue the objective assimilation of a forecast and observations into a nowcast frontal path using the optimum interpolation-based PATHFINDER. As part of this effort, determine the relative weights of the observations, which have a range of ages, and the forecast. Also, reexamine optimum correlation scales and the phase speeds to be used in PATHFINDER (this issue will be investigated more closely, once some operational experience with the model has been gained). Add more user-friendly graphics and menus to simplify gross error checks and other operator interactions with the system.

Determine how complementary, rather than competing, products can be obtained using the DART system in the NAVOCEANO environment, which has a skilled operator available, and the automated FNOC environment. Develop operating procedures for guiding the operator in making the maximum value-added contribution to the product.

Develop and transition a data assimilation technique that assimilates a history of incomplete frontal paths into a single forecast of the frontal path.

Develop a technique to objectively determine the north wall of the Gulf Stream from the model's surface topography field. Consider other techniques, such as inferring the frontal axis from the model transport or velocity field.

Implement the software to run the model on the Class VII computer in an automated mode with multiple initializations. Use the envelope of the resulting forecasts to assign an uncertainty to the model forecast for use in PATHFINDER.

Establish guidance for the operator to use the distribution of data gaps and ages in setting up a series of sensitivity experiments that are specific to that distribution.

XI. References

Bretherton, F. P., R. E. Davis, and C. B. Fandry (1976). A technique for objective analysis and design of oceanographic experiments applied to MODE-73. *Deep-Sea Research* 23:559–582.

Clancy, R. M., P. A. Phoebus, and K. D. Pollak (1990). An operational global-scale ocean thermal analysis system. *Journal of Atmospheric and Oceanic Technology* 7:233–254.

Fox, D. N., M. R. Carnes, and J. L. Mitchell. Circulation Model Experiments of the Gulf Stream Using Satellite-Derived Fields. Naval Oceanographic and Atmospheric Research Laboratory, Stennis Space Center, MS, submitted.

Fox, D. N., H. E. Hurlburt, J. D. Thompson, Z. R. Hallock, and G. A. Ransford (1988). Oceanic data assimilation and prediction using remotely sensed data. *Proceedings of the Pacific Congress on Marine Science and Technology (PACON 88)*, Honolulu, HI, pp. OST2/17–23.

Hurlburt, H. E., D. N. Fox, and E. J. Metzger (1990). Statistical inference of weakly correlated subthermocline fields from satellite altimeter data. *Journal of Geophysical Research* 95:11375–11409.

Hurlburt, H. E. and J. D. Thompson (1980). A numerical study of loop current intrusions and eddy shedding. *Journal of Physical Oceanography* 10:1611–1651.

Thompson, J. D. and W. J. Schmitz, Jr. (1989). A limited-area model of the gulf stream: design, initial experiments, and model-data intercomparison. *Journal of Physical Oceanography* 19:791–814.

Wallcraft, A. J. (1990). The NOARL Layered Ocean Model User's Guide. Naval Oceanographic and Atmospheric Research Laboratory, Stennis Space Center, MS, NOARL Report 35.

Appendix

DART Transition Plan

SER320/74
Feb 12, 1990

FROM: Commanding Officer, NOARL

TO: Office of Chief of Naval Research (Code 120M)
Commander, Naval Oceanography Command
Commanding Officer, Fleet Numerical Oceanography Command
Commanding Officer, Naval Oceanographic Office

SUBJ: DART Transition Plan; promulgation of

REF: (a) Memorandum of Understanding for Ocean-Atmospheric-Acoustic Model Transition
(b) DART Transition Meeting 8-9 Jan 90 at NOARL
(c) DART Transition Workshop 31 Jan-2 Feb 90 at NOARL (West)

ENCL: (1) DART Transition Plan

1. In accordance with ref (a), meetings were held, ref (b) and (c), to develop the DART Transition Plan. During the meeting on 8-9 Jan 90 (ref b) programmatic guidance from the RDT & E program manager was provided for the transition of the DART system, first to NAVOCEANO and then to FLENUMOCEANCEN. The meeting at NOARL (West) between NOARL, NAVOCEANO, and FLENUMOCEANCEN, ref (c), was conducted to refine the draft transition plan.

2. The DART Transition Plan, enclosure (1), is hereby promulgated for programmatic coordination and appropriate action.

1. PROGRAM TITLE: Navy Ocean Modeling and Prediction Program

2. PROGRAM ELEMENT: 63207N

3. PROJECT TITLE: Data Assimilation Research and Transition

4. POINTS OF CONTACT:

NOARL:

Dr. Joseph W. McCaffrey, Jr., Division Head	(601) 688-4870
Dr. Jim L. Mitchell, DART Manager	(601) 688-4813
Dr. Theodore J. Bennett, Jr., Transition Lead	(601) 688-4704

FNOC:

Mr. R. Michael Clancy	(408) 647-4414

NAVOCEANO:

Mr. Andrew A. Johnson	(601) 688-4403

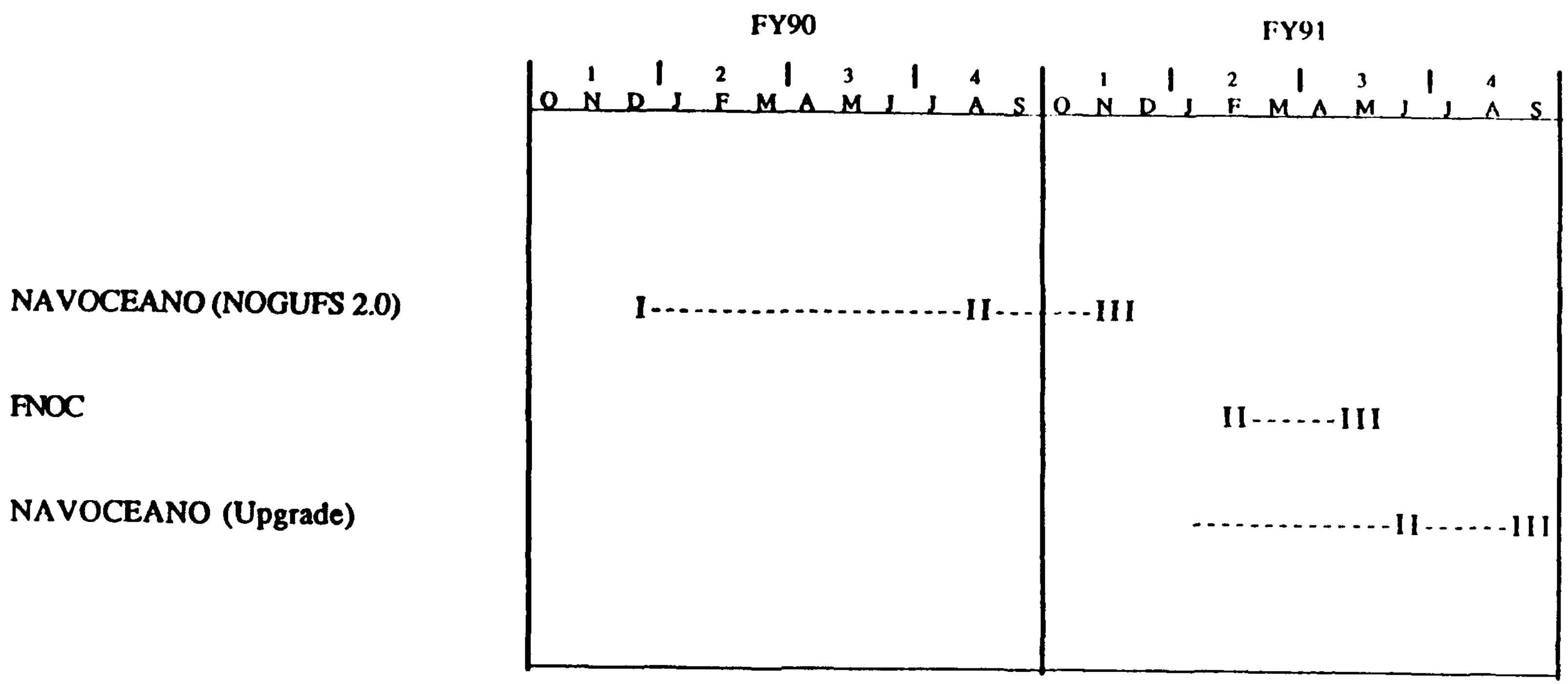

Figure 1. Gulf Stream forecast model delivery schedule.

I. PRODUCT DESCRIPTION

The objective of the Data Assimilation Research and Transition (DART) project at the Naval Oceanographic and Atmospheric Research Laboratory (NOARL) is to develop and transition upgrades to the Navy's capability to monitor and describe the ocean environment. Ongoing work at both the U.S. Naval Oceanographic Office (NAVOCEANO) and the Fleet Numerical Oceanography Center (FNOC) are effectively coupled to the DART work. This objective is particularly important because the data sparsity of the ocean, particularly at depth, is a major impediment to the effective use of the Navy's capability.

This document is a Transition Plan that describes NOARL's plans to transition a system for making fourteen day forecasts of the Gulf Stream frontal path. The initial state for the forecast is generated by the application of Optimum Thermal Interpolation System (OTIS) to the initial Bogus, which is a map of front and ring positions together with estimates of their various parameters. Operationally, the forecast frontal position will be used to fill in data gaps during the construction of the Bogus.

In addition, the Plan follows the draft Memorandum of Understanding between NOARL, NAVOCEANO and FNOC that sets forth a generally agreed upon process for the transition of oceanic, atmospheric and acoustic models to operational Navy use.

Other products, their upgrades and their application to new regions of the ocean are anticipated during the outyears and will be described in subsequent transition plans. For example, a plan for the transition of software for generating synthetic profiles of temperature, salinity and sound speed from satellite altimeter-derived sea surface topography has been presented by NAVOCEANO (NAVOCEANO ltr Ser OS/9049 of 19 DEC 88) and accepted by NOARL, then the Naval Ocean Research and Development Activity (NORDA ltr Ser 320/83 of 21 MAR 89). This work will be pursued in FY 91.

A. System Description

This software is designed to forecast the time evolution of the path of the Gulf Stream and the drift of its associated rings. The Plan calls for an immediate transition of the software to NAVOCEANO followed by a transition to FNOC. If the transition to NAVOCEANO is successful, then the model will be designated the NOGUFS 2.0 model.

This model consists of several groups of software modules. These groups are (1) OTIS, a thermal analysis system developed at FNOC; (2) initialization software that constructs the fields required by the circulation module from the thermal analysis; and (3) a two-active-layer, primitive equation circulation model. All of this software runs on a Hewlett-Packard (HP) 9000 Series 835 computer.

The input to the system is the Bogus prepared operationally by NAVOCEANO on a HP 9000 Series 835 computer. A thermal analysis of the model domain is generated by applying OTIS to the Bogus. In its initial implementation, no data other than the Bogus is assimilated by OTIS. The fields required by the circulation module are generated from the thermal analysis by the initialization software. Various parameters for the modules, such as the correlation scales used in OTIS and the eddy viscosity used in the circulation module, are set by the developer or modified as part of an upgrade; the parameters are not normally changed day-to-day by the operator.

Output will be two-week forecasts of front and ring positions. These output maps will be used to interpolate the path of the Gulf Stream across data gaps, i.e. across cloudy areas where infrared (IR) data are unavailable.

The operator can make a significant contribution to the forecast. Although operator involvement can be limited to initiating the job streams and monitoring the results (check for job failures, etc.), the value added by the operator is his/her capability to make a series of sensitivity experiments. For example, if an operator questions a forecast of the frontal path across a data gap, then the operator could potentially examine the data that were used to construct the initial Bogus for the forecast. Using the operator's judgement, a series of sensitivity tests could be made by forecasting from a modified initial Bogus. These sensitivity tests would provide additional guidance to the operator that constructs the BOGUS.

During the OPTEST phase at NAVOCEANO, the Gulf Stream forecast model will be transitioned to FNOC, where the emphasis is on coupling the mesoscale oceanography to the large-scale circulation of the ocean. This coupling will ultimately be realized by either a single high-resolution basin model or a nesting of the DART model within a more-coarse resolution basin model. In addition, the DART model will provide important forcing data to the other components of the FNOC ocean nowcast and forecast system. The transition of the DART model to FNOC is a first step toward implementing this capability and developing a base of experience.

At FNOC, the model will be required to operate in a fully-automated, "hands-off" mode. In addition, it will be required to interface with the already operational Gulf Stream OTIS and Thermodynamic Ocean Prediction System (TOPS) models at FNOC. In particular, the model will be required to run on the same model domain as the OTIS and TOPS models.

B. Financial Summary

See attached sheet.

C. Required Technical Performance

TRANSITION BUDGET

FY	Program Element	Cost ($K)
1990	63207N	212
1991	63207N	100
	Total	312

Table 1. Summary of 6.3 funding required for the transition of the DART Gulf Stream forccast model.

The technical objective of this work is to make one- and two-week forecast of the Gulf Stream frontal position that is better than persistence. A persistence forecast assumes that the frontal position does not vary from its initial path. Forecast quality is determined by comparing a model forecast to the validation data on the stream position for the time that the forecast is valid. The average absolute difference between the forecast frontal position and the validation data along the path segments where such data are available is the objective measure used to determine quality. The same measure is applied to the persistence forecast. Typical performance is estimated by averaging over a number of cases.

Several criteria are used to judge the model's performance. Persistence can be a very good forecast in a situation where no events such as ring break-off or coalescence occur. On the other hand, persistence can be a poor forecast in situations where an event occurs or a meander has a significant change in shape or position. In both event and non-event cases, the model is expected to typically make a better forecast than persistence.

An important caveat must be stated. The operational Bogus is determined primarily from infrared (IR) data. Under cloudy conditions, there are no data to determine the frontal position and hence the initial state of the model. The sensitivity of the model forecast to data gaps in the Bogus is an issue that will be studied during this transition. However, it is clear that there will be times when there are insufficient data, i.e. too many data gaps, to determine the initial state.

D. Required Operational Characteristics

The initial transition is to NAVOCEANO, where the model will be required to run on a HP 9000 Series 835 computer. The technical knowledge and judgement required to operate the model must be within the capabilities of the personnel of the NAVOCEANO Operational Oceanography Center (OOC) that will be asked to run the model. Using operationally available data, the model will be required to make one- and two-week forecasts of the Gulf Stream path that are statistically better than persistence and climatology.

Following the transition to NAVOCEANO, the Gulf Stream forecast model will be implemented on the Cyber 205 supercomputer at FNOC. At FNOC, the model will be required to operate in a fully-automated, "hands-off" mode. In addition, it will be required to interface with the already operational Gulf Stream OTIS and Thermodynamic Ocean Prediction System (TOPS) models at FNOC. In particular, the model will be required to run on the same model domain as the OTIS and TOPS models.

E. Critical Technical Issues

The Proof-of-Concept results presented by NOARL to the Commander, Naval Oceanography Command (COMNAVOCEANCOM) Independent Model Review Panel (CIMREP) in December 1989 demonstrate that the DART model has a skill at forecasting the Gulf Stream path that is significantly better than persistence at one and two weeks when there are abundant data available for initialization. Four critical issues are unanswered by those results. These issues are:

(1) what is the sensitivity of the forecast skill to spatial gaps in the data available for initialization?

(2) what skill does the model have in an operational environment given the sensitivity of the model to the initial state and the limited data typically available for initialization? This question can also be posed as what fraction of the time does the available data justify using the model?

(3) how are the model forecast of frontal position as well as the available data objectively assimilated into a single Bogus message. What is the value added by the forecast to the Bogus? and

(4) the most effective use of this system requires an operator to interact with the system. For example, if an operator questions a forecast of the frontal path across a data gap, then the operator could potentially examine the data that were used to construct the initial Bogus for the forecast. Using the operator's judgement, a series of sensitivity tests could be made by forecasting from a modified initial Bogus. However, the skills necessary for this interaction and the proper guidelines for the operator have not been worked out.

II. MILESTONES

A Technical Validation Panel has been named by the Commanding Officer of NOARL to monitor the progress of the transition and offer technical advice to the transition lead. The chair of the Panel is from NOARL. Representatives from NAVOCEANO, FNOC and outside experts are also on the Panel. A key objective of the Panel is to facilitate the building of a consensus between NOARL and the operational commands on just what the products will be, how they will be used, what technical approach will be used to transition these products and what will the milestones be.

The Panel will also assist in assigning responsibility for each milestone. The general guidelines are that NOARL will have the lead responsibility for validating the model and preparing the documentation. NOARL will also provide technical assistance during the Operation Evaluation (OPEVAL) phase. NAVOCEANO and FNOC will have the lead during the OPEVAL phase. In addition, FNOC will have the lead, with assistance from NOARL and NAVOCEANO, in implementing the software on the Cyber 205 computer at FNOC.

Figure 2 is a milestone chart for FY 90 that breaks the transition work down into fifteen tasks. These tasks are:

(1) Document the Proof-of-Concept experiments. This 6.2 work will be the written basis for the decision to proceed with the transition.

(2) Name the Technical Validation Panel.

(3) Prepare a transition plan.

(4) Estimate the value added by the forecast to the Bogus and the forecast sensitivity to data gaps.

(5) Estimate forecast performance.using operational Boguses.

(6) Prepare research quality Bogus messages for several weeks of very high quality IR data from 1989 that have recently become available.

(7) Make new tests using several weeks of very high quality IR data from 1989 that have recently become available.

(8) Finalize software. The software, especially the input and output data flows and job control commands, must be put into a user-friendly form and frozen so that documentation can begin.

(9) Prepare graphics software.

(10) Document the model. Documentation of the model is required at the time that the model enters Operational Check (OPCHECK) at NAVOCEANO. This documentation includes a User's Guide, a Software Requirement Specification (SRS), a Software Test Description (STD) and a Software Design Document (SDD). The documentation shall conform to the NAVOCEANO standards for the documentation of environmental systems and products. These standards are defined in "Software Documentation Standards and Coding Requirements for Environmental System Product Development (September 1988). NAVOCEANO is currently revising these standards. A draft revision of these standards is attached.

(11) Write the Validation Test Report (VTR). The VTR presents the technical results of the test and evaluation work and provides guidance to the user on the appropriate use of the system. A quick look document will be delivered to NAVOCEANO at the start of the OPCHECK phase with a more thorough report being prepared during the following months.

(12) NOARL Certification. The Commanding Officer of NOARL or his designated representative will certify to NAVOCEANO that the DART Gulf Stream forecast model is ready for OPCHECK at NAVOCEANO.

(13) Operational Check (OPCHECK).

(14) Operational Test (OPTEST).

(15) Monitor OPEVAL. NOARL will have the opportunity to observe the OPEVAL phase of the transition, identify problems that arise and feedback to the model developers information on the model's performance and problems.

Figure 3 is a milestone chart for FY 91 that breaks the transition work down into nine tasks. These tasks are:

(1) Continuation of the OPTEST. The OPTEST at NAVOCEANO will continue into FY 91.

(2) Continuation of the preparation of the VTR. The preparation of the VTR will continue into FY 91.

(3) Implementation at FNOC.

(4) OPCHECK at FNOC.

(5) OPTEST at FNOC.

(6) Monitor operational use of model. The objective of this task is to participate in the ongoing evaluation of the model and to promote the flow of information between the operational commands and the research team. The experience of the operational commands will identify to the research team the actual performance of the model as well as the most critical issues.

(7) Implement upgrade at NAVOCEANO. The NOGUFS 2.0 system will be upgraded via the implementation of the OTIS 3.0 system and the GDEM dynamic climatology.

(8) Upgrade OPCHECK.

(9) Upgrade OPTEST.

III. TEST AND EVALUATION

The test and evaluation of the Gulf Stream forecast model have six thrusts. These thrusts are:

(1) Forecast sensitivity to data gaps. Realistically, spatial gaps can be expected in the data used to determine the Bogus. The sensitivity of the model forecast to these gaps will be examined using the data sets that were used to make the test cases that were presented to the CIMREP panel. Using the abundant IR data in these data sets, virtually the entire initial and verification Gulf Stream paths can be mapped. In the sensitivity tests, subsections of the initial path will be deleted and a variety of interpolations made across the resulting gap. A forecast will then be made that can be compared to the forecast based on gapless data as well as persistence. In order to objectively examine the sensitivity, the average absolute distance between the forecast path and the verification data is estimated. The actual difference between the forecast path and the verification data as a function of path length will also be examined.

(2) Real-time forecast sensitivity studies. Real-time sensitivity tests will be used to examine how well the forecast system can be expected to perform in an operational environment. Each week the operational Bogus will be used to initialize the forecast system. The average absolute difference between the forecast path and the verification data will be determined for the subsections of the path that have verification data. How well the model forecast matches the subsections that have data is a measure of how well the forecast can interpolate across the data gaps. Systematic problems with the model as well as the sparsity of the initialization data contribute to the forecast error. For what fraction of the time will there be adequate data to initialize the model and make a skillful forecast?

(3) Value added. The data sets that were used to generate the results presented to the CIMREP panel will also be used to estimate the value added by the forecast to the Bogus. A Gulf Stream Bogus obtained from these data will have subsections of the Stream path deleted. A two-week forecast valid at the time of the Bogus will be used by the NAVOCEANO Pathfinder system to interpolate across the data gaps. Pathfinder is an optimum interpolation-based system that assumes a first guess of the path, say the forecast, and then assimilates the data (subsections of path verification data) into it. Does the forecast add any value to the Bogus across the data gaps?

(4) New tests using abundant data. Several, multi-week periods of excellent IR data from 1989 have recently become available. Research-grade Boguses will be made for these periods and a series of forecast experiments performed. These experiments will supplement the results presented to the CIMREP panel.

(5) Sensitivity of the model initialization to the version of OTIS. The OTIS 2.1 system that was used to generate the results presented to the CIMREP panel is expected to be replaced by the OTIS 3.0 system in a major upgrade scheduled for FY 91. Although the DART team does not expect the conclusions based on OTIS 2.1 to change when OTIS 3.0 replaces OTIS 2.1, a series of side-by-side experiments will be performed in order to verify this expectation.

(6) Other measures of forecast performance. Measures other than the average absolute distance between paths must be developed in order to more fully understand and grade the quality of the forecasts. These measures ought to examine potential systematic problems as phase shifts and over- or under-amplification of meanders.

Gulf Stream Forecast Model

(1) NAVOCEANO OPTEST

(2) VTR

(3) FNOC Implementation

(4) FNOC OPCHECK

(5) FNOC OPTEST

(6) Monitor/Feedback

(7) Implement NOGUFS 2.1

(8) NOGUFS 2.1 OPCHECK

(9) NOGUFS 2.1 OPTEST

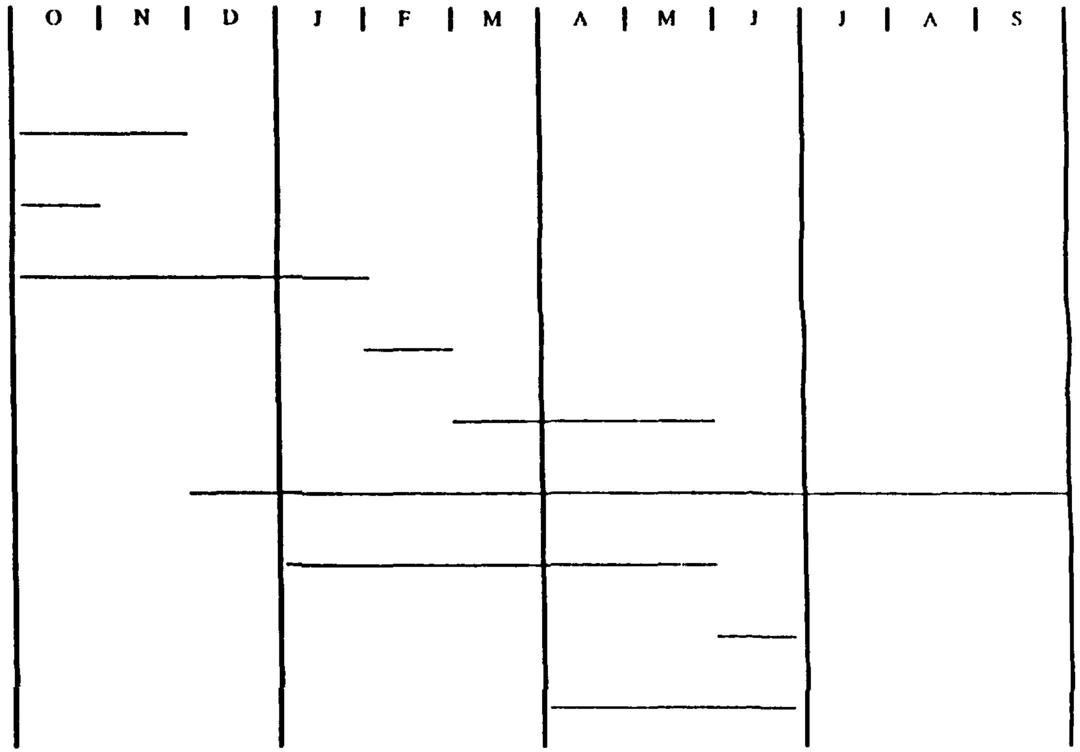

Figure 3. FY91 Milestones

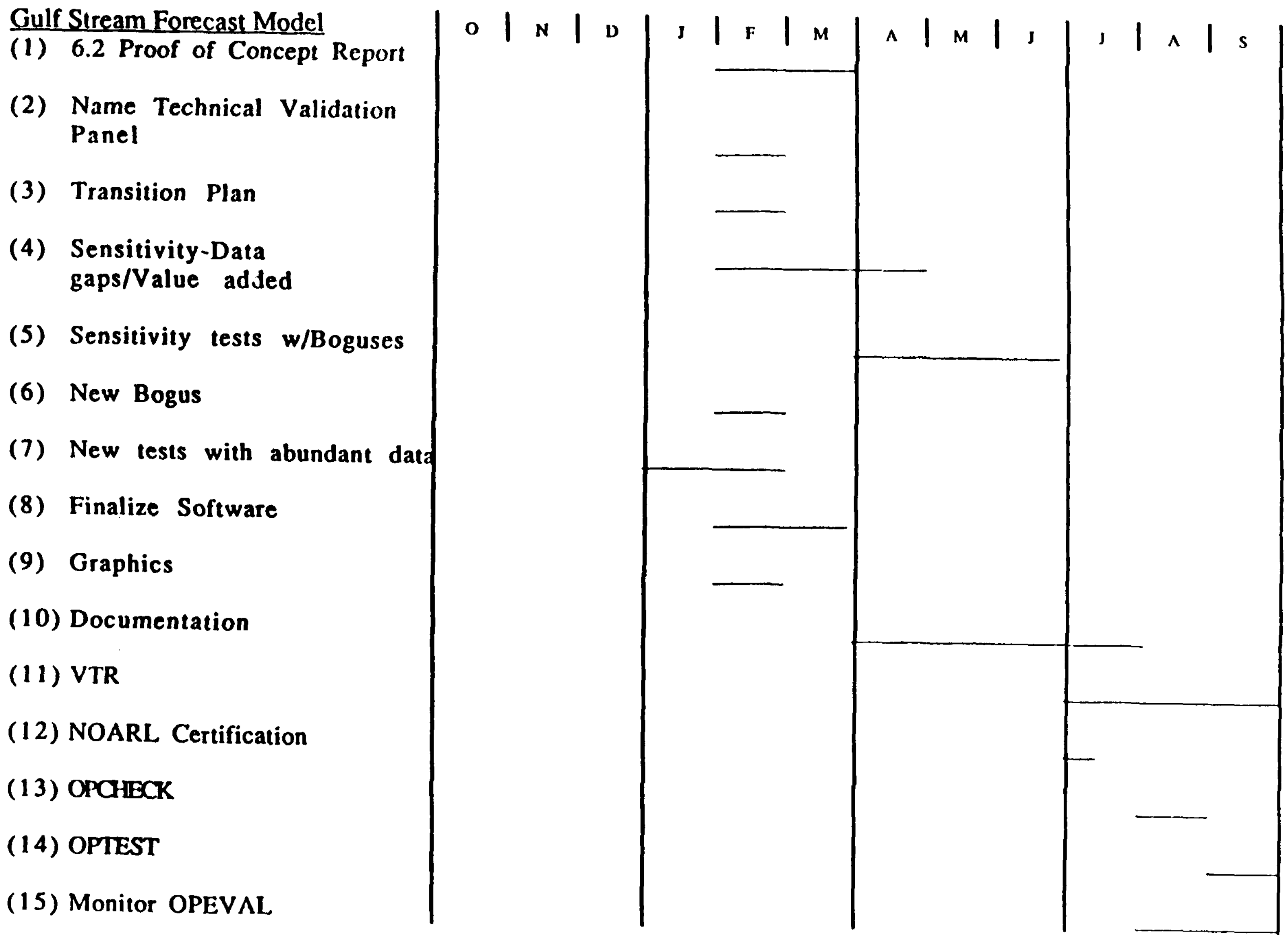

Figure 2. FY90 Milestones

Distribution List

Applied Physics Laboratory
Johns Hopkins University
Johns Hopkins Road
Laurel MD 20707

Applied Physics Laboratory
University of Washington
1013 NE 40th St.
Seattle WA 98105

Applied Research Laboratory
Pennsylvania State University
P.O. Box 30
State College PA 16801-0030

Applied Research Laboratories
University of Texas at Austin
P.O. Box 8029
Austin TX 78713-8029

Assistant Secretary of the Navy
Research, Development & Acquisition
Navy Department
Washington DC 20350-1000

Chief of Naval Operations
Navy Department
Washington DC 20350-2000
Attn: OP-71
OP-987

Chief of Naval Operations
Oceanographer of the Navy
U.S. Naval Observatory
34th & Massachusetts Ave. NW
Washington DC 20392-1800
Attn: OP-096
OP-0961B

David W. Taylor Naval Research Center
Bethesda MD 20084-5000
Attn: Commander

Defense Mapping Agency
Systems Center
8613 Lee Hwy.
Mail Stop A-13
Fairfax VA 22031-2138
Attn: Code PRN

Fleet Antisub Warfare Tng Ctr-Atl
Naval Station
Norfolk VA 23511-6495
Attn: Commanding Officer

Fleet Numerical Oceanography Center
Monterey CA 93943-5005
Attn: Commanding Officer

National Ocean Data Center
1825 Connecticut Ave., NW
Universal Bldg. South, Rm. 206
Washington DC 20235

Naval Air Development Center
Warminster PA 18974-5000
Attn: Commander

Naval Air Systems Command HQ
Washington DC 20361-0001
Attn: Commander

Naval Civil Engineering Laboratory
Port Hueneme CA 93043
Attn: Commanding Officer

Naval Coastal Systems Center
Panama City FL 32407-5000
Attn: Commanding Officer

Naval Facilities Engineering
Command HQ
200 Stovall St.
Alexandria VA 22332-2300
Attn: Commander

Naval Oceanographic Office
Stennis Space Center MS 39522-5001
Attn: Commanding Officer
Library
Code TD, L. Bernard

Naval Oceanography Command
Stennis Space Center MS 39529-5000
Attn: Commander

Naval Oceanographic & Atmospheric
Research Laboratory
Atmospheric Directorate
Monterey CA 93943-5006
Attn: Code 400

Naval Oceanographic & Atmospheric
Research Laboratory
Stennis Space Center MS 39529-5004
Attn: Code 100
Code 105
Code 115
Code 125L (10)
Code 125P
Code 200
Code 300

Naval Ocean Systems Center
San Diego CA 92152-5000
Attn: Commander

Naval Postgraduate School
Monterey CA 93943
Attn: Superintendent

Naval Research Laboratory
Washington DC 20375
Attn: Commanding Officer
Library, M. Bradley (3)

Naval Sea Systems Command HQ
Washington DC 20362-5101
Attn: Commander

Naval Surface Warfare Center Det
Silver Spring
White Oak Laboratory
10901 New Hampshire Ave.
Silver Spring MD 20903-5000
Attn: Officer in Charge
Library

Naval Surface Warfare Center
Dahlgren VA 22448-5000
Attn: Commander

Naval Underwater Systems Center
Newport RI 02841-5047
Attn: Commander

Naval Underwater Systems Center Det
New London Laboratory
New London CT 06320
Attn: Officer in Charge

Office of Naval Research
800 N. Quincy St.
Arlington VA 22217-5000
Attn: Code 10D/10P, Dr. E. Silva
Code 112, Dr. E. Hartwig
Code 12
Code 10

Office of Naval Research
ONR European Office
PSC 802 Box 39
FPO AE 09499-0700
Attn: Commanding Officer

Office of Naval Technology
800 N. Quincy St.
Arlington VA 22217-5000
Attn: Code 20, Dr. P. Selwyn
Code 228, Dr. M. Briscoe
Code 228, CDR. L. Bounds
Code 22, Dr. T. Warfield

Scripps Institution of Oceanography
University of California
291 Rosecrans St.
San Diego CA 92106-3505

Space & Naval Warfare Sys Com
Director of Navy Laboratories
SPAWAR 005
Washington DC 20363-5100
Attn: Commander

Woods Hole Oceanographic Institution
P.O. Box 32
Woods Hole MA 02543
Attn: Director

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. Agency Use Only *(Leave blank).*

2. Report Date. April 1992

3. Report Type and Dates Covered. Final

4. Title and Subtitle.

Validation Test Report for the First-Generation Dart Gulf Stream Forecasting System

5. Funding Numbers.

Work Unit No.	93200B
Program Element No	0603207N
Project No	X2008
Task No.	
Accession No.	DN259001

6. Author(s).

Theodore J. Bennett, Jr. and *Robert C. Rhodes

7. Performing Organization Name(s) and Address(es).

Naval Oceanographic and Atmospheric Research Laboratory
Ocean Science Directorate
Stennis Space Center, Mississippi 39529-5004

8. Performing Organization Report Number.

NOARL Report 20

9. Sponsoring/Monitoring Agency Name(s) and Address(es).

Space and Naval Warfare Systems Command
PDW-141
Washington, DC 20363-5100

10. Sponsoring/Monitoring Agency Report Number.

11. Supplementary Notes.

Naval Oceanographic Office, Stennis Space Center, Mississippi 39529

12a. Distribution/Availability Statement.

Approved for public release; distribution is unlimited. Naval Oceanographic and Atmospheric Research Laboratory, Stennis Space Center, Mississippi 39529-5004.

12b. Distribution Code.

13. Abstract *(Maximum 200 words).*

Acoustic propagation within the Northwest Atlantic is highly dependent on the meandering Gulf Stream front and its associated drifting rings, the socalled mesoscale structure of the ocean. Considerable variability in space and time characterize this structure.

This report documents the validation of a system that makes nowcasts and forecasts of the Gulf Stream frontal path. The validation was done as part of the transition of the first-generation Data Assimilation Research and Transition (DART) model to operational Navy use. Two sets of experiments designed to evaluate the forecast skill of the system in a quasi-operational environment are the crux of the validation. In one set of experiments, operational front and ring maps are used to initialize and verify 1- and 2week forecast experiments. In the other set of experiments, the value the DART model adds to the present operational capability is estimated. The proposed operational use of the model is simulated by setting up a simple data assimilation system that uses a 1-week forecast to interpolate across long data gaps that arise during the mapping of front and ring positions. The gappy paths are constructed by superimposing observed data gaps on complete frontal paths. This report also provides guidance to an operator on how to use a model forecast to fill in the data gaps.

Both sets of experiments show the system to be a statistically significant improvement over an assumption of persistence, i.e., no change over time. The principal technical issues identified in the report are the inference of the Gulf Stream north wall from the forecasted axis path, as well as the need to develop and test a more objective means to assimilate observations and a forecast into a composite frontal path. It is recommended that these issues be pursued.

14. Subject Terms.

ocean models, data assimilation

15. Number of Pages. 43

16. Price Code.

17. Security Classification of Report.	18. Security Classification of This Page.	19. Security Classification of Abstract.	20. Limitation of Abstract.
Unclassified	Unclassified	Unclassified	Same as report